No Eigenvalues Outside the Support of the Limiting Spectral Distribution of Information-Plus-Noise Type Matrices

Zhidong Bai[†1] and Jack W. Silverstein[‡2]

†*KLAS MOE & School of Mathematics and Statistics, Northeast Normal University, Changchun, Jilin 130024, China;*
Email : baizd@nenu.edu.cn

‡*Department of Mathematics, Box 8205, North Carolina State University, Raleigh, NC 27695-8205, USA;*
Email : jack@math.ncsu.edu

Abstract

We consider a class of matrices of the form $\mathbf{C}_n = (1/N)(\mathbf{R}_n+\sigma\mathbf{X}_n)(\mathbf{R}_n+\sigma\mathbf{X}_n)^*$, where $\mathbf{X}_n$ is an $n \times N$ matrix consisting of independent standardized complex entries, $\mathbf{R}_j$ is an $n\times N$ nonrandom matrix, and $\sigma > 0$. Among several applications, $\mathbf{C}_n$ can be viewed as a sample correlation matrix, where information is contained in $(1/N)\mathbf{R}_n\mathbf{R}_n^*$, but each column of $\mathbf{R}_n$ is contaminated by noise. As $n \to \infty$, if $n/N \to c > 0$, and the empirical distribution of the eigenvalues of $(1/N)\mathbf{R}_n\mathbf{R}_n^*$ converge to a proper probability distribution, then the empirical distribution of the eigenvalues of $\mathbf{C}_n$ converges a.s. to a nonrandom limit. In this paper we show that, under certain conditions on $\mathbf{R}_n$, for any closed interval in $\mathbb{R}^+$ outside the support of the limiting distribution, then, almost surely, no eigenvalues of $\mathbf{C}_n$ will appear in this interval for all n large.

Keywords : Empirical spectral distribution, Stieltjes transform, information-plus-noise model.

[1]Bai was supported by the NSF China, Grant 10871036 and the NUS Grants R-155-000-079-112 and R-155-000-096-646.

[2]Silverstein's work is supported by U.S. Army Research Office under Grant W911NF-09-1-0266

1 Introduction.

For $n = 1, 2, \ldots$ let $\mathbf{X}_n = (X_{ij})$ be an $n \times N$ matrix containing independent standardized complex entries ($\mathrm{E}X_{ij} = 0$, $\mathrm{E}|X_{ij}|^2 = 1$), and let $\mathbf{R}_n$ be an $n \times N$ nonrandom matrix. The matrix

$$\mathbf{C}_n = \frac{1}{N}(\mathbf{R}_n + \sigma\mathbf{X}_n)(\mathbf{R}_n + \sigma\mathbf{X}_n)^*$$

can be viewed as a sample correlation matrix formed from the N sampled vectors $\mathbf{r}_j + \mathbf{x}_j$, where $\mathbf{r}_j$ and $\mathbf{x}_j$ are the j-th columns of $\mathbf{R}_n$ and $\mathbf{X}_n$, respectively. Among several applications for the use of $\mathbf{C}_n$ is the case where information is contained in the matrix $(1/N)\mathbf{R}_n\mathbf{R}_n^*$, but each $\mathbf{r}_j$ is contaminated by additive noise. For this reason $\mathbf{C}_n$ has been described as the information-plus-noise type matrix. A result on the behavior of the eigenvalues of $\mathbf{C}_n$ have been obtained in Dozier and Silverstein (2007a), under the assumption the entries of $\mathbf{X}_n$ are i.i.d. The result is expressed in terms of the empirical spectral distribution function $F^{\mathbf{C}_n}$ (where, for any square matrix $\mathbf{A}$ with only real eigenvalues, $F^{\mathbf{A}}(x)$ is the proportion of eigenvalues of $\mathbf{A} \leq x$). It is shown that, if $N = N(n)$ and $c_n \equiv n/N \to c > 0$ and $H_n \equiv F^{(1/N)\mathbf{R}_n\mathbf{R}_n^*}$ converges in distribution to H, a proper probability distribution function (p.d.f.), as $n \to \infty$, then, with probability 1, $F^{\mathbf{C}_n}$ converges in distribution to F, a p.d.f. which is nonrandom. The proof uses the Stieltjes transform method, where, for any p.d.f. G, the Stieltjes transform of G is defined as

$$m_G(z) \equiv \int \frac{1}{\lambda - z} dG(\lambda), \qquad z \in \mathbb{C}^+ \equiv \{z \in :\Im z > 0\}.$$

Due to the inversion formula

$$G([a,b]) = \frac{1}{\pi} \lim_{\eta \to 0} \int_a^b \Im m_G(\xi + i\eta) d\xi$$

(a, b continuity points of G), weak convergence of p.d.f.'s is equivalent to convergence of Stieltjes transforms.

It is shown in Dozier and Silverstein (2007a) that, for any $z \in \mathbb{C}^+$, with probability 1, $m_{F^{\mathbf{C}_n}}(z)$ converges to $m^0 = m_F(z)$, which satisfies the equation

$$m = \int \frac{1}{\frac{t}{1+\sigma^2 cm} - (1 + \sigma^2 cm) + \sigma^2(1-c)} dH(t). \tag{1.1}$$

It is also proven that m is the only solution to (1.1) for which $m \in \mathbb{C}^+$ and $\Im zm \geq 0$ (note: it is straightforward to show that for any p.d.f. G which has mass concentrating on the nonnegative reals, and for $z \in \mathbb{C}^+$, we have $m_G(z) \in \mathbb{C}^+$ and $zm_G(z) \in \mathbb{C}^+$).

The purpose of this paper is to prove, along the same lines as in Bai and Silverstein (1998) and Paul and Silverstein (2007), the nonexistence of eigenvalues outside the support of F. It is necessary at this point to review some of the properties of F and m_F, obtained in Dozier and Silverstein (2007b). It is shown that for all $x \in \mathbb{R} - \{0\}$, $\lim_{z \in \mathbb{C}^+ \to x} m_F(z) \equiv m^0(x)$ exists. From this it follows that F has a continuous derivative f on $\mathbb{R} - \{0\}$ given by $f(x) = \frac{1}{\pi}\Im m^0(x)$. Along the way in proving this result it was shown that

$$|m_F(z)| \leq \left(\frac{1}{\sigma^2 c|z|}\right)^{1/2}, \quad \text{for} \quad z \in \mathbb{C}^+. \tag{1.2}$$

Let $b = 1 + \sigma^2 c m^0$. Then, with

$$w(z) \equiv b^2(z)z - b(z)\sigma^2(1-c), \tag{1.3}$$

(1.1) can be written as

$$\frac{1}{\sigma^2 c}\left(1 - \frac{1}{b(z)}\right) = m_H(b^2(z)z - b(z)\sigma^2(1-c)) = m_H(w(z)). \tag{1.4}$$

Let S_G denote the support of the p.d.f. G. It is shown in Dozier and Silverstein (2007b) that

$$x \neq 0 \text{ and } x_0 \in S_F^c \Longrightarrow w(x_0) \in S_H^c. \tag{1.5}$$

Moreover, for all $z \in (\mathbb{C}^+ \cup \mathbb{R}) - \{0\}$, $\Re b > 0$. We see then, from continuity, that (1.4) holds for $x \in \mathbb{R} - \{0\}$ in a neighborhood of x_0. We will use the fact that, on intervals outside the support of p.d.f. G, $m_G(x)$ exists, and is increasing, which implies its inverse exists, and is increasing. Therefore, for any interval $I \in \mathbb{R} - \{0\}$ contained in S_F^c, there exists an interval $J \in \mathbb{R}^+$ for which we have the inverse of m_F expressed in terms of b:

$$x(b) = \frac{1}{b^2} m_H^{-1}\left(\frac{1}{\sigma^2}c\left(1 - \frac{1}{b}\right)\right) + \frac{1}{b}\sigma^2(1-c) \quad b \in J. \tag{1.6}$$

Moreover, we have $x'(b) > 0$ for $b \in J$.

Conversely, it is shown that, for any interval $L \subset S_H^c$, there is an interval $J \subset \mathbb{R}^+$ for which the mapping $b \to \frac{1}{\sigma^2 c}(1 - 1/b)$ takes J into the range of m_H applied to L, so that the function in (1.6) is well defined. Moreover, if $x'(b) > 0$ then $x(b) \in S_F^c$ and $b = 1 + \sigma^2 c m_F(x(b))$.

We concentrate on an interval $[a_1, a_2]$ for which $a_1 > 0$ and $[a_1, a_2] \subset (b_1, b_2)$, where $b_1 > 0$ and $(b_1, b_2) \subset S_F^c$. From (1.5) and the fact that $w(\cdot)$ is analytic in S_F^c, we see that $w((b_1, b_2)) = (c_1, c_2) \subset S_H^c$. In order to make any conclusion on the nonappearance of eigenvalues of $\mathbf{C}_n$ in (b_1, b_2), this interval along with (c_1, c_2) must somehow require assumptions on c_n and the eigenvalues of $(1/N)\mathbf{R}_n\mathbf{R}_n^*$, for n large. This is done by utilizing equation (1.1) for each n. Let m_n^0 and b_n^0 be the respective solutions to (1.1) and (1.4) with H and c replaced by H_n and c_n respectively, let w_n^0 be defined by (1.3) in terms of b_n^0, and let F^{c_n,H_n} denote the corresponding p.d.f. We impose the following condition:

$$\text{For all } n \text{ large, } (b_1, b_2) \subset S^c_{F^{c_n,H_n}}. \tag{1.7}$$

Because of (1.2) and the uniqueness of solutions to (1.1), it follows that $m_n^0(z) \to m^0(z)$ as $n \to \infty$ for all z with $\Im z \neq 0$. Since the m_n^0 are analytic in a neighborhood of (b_1, b_2) in $\mathbb{C}$, we have by Vitali's convergence theorem (Titchmarsh (1939), p. 168), $m_n^0(x) \to m^0(x)$ as $n \to \infty$ for all $\in (b_1, b_2)$. Upon differentiation of the extreme sides of (1.4) with respect to $x \in (b_1, b_2)$, we find that $w'(x) > 0$. Therefore

$$w(b_1) < w(a_1) \quad \text{and} \quad w(a_2) < w(b_2).$$

Hence, there exists an $\epsilon >$, such that, for all n large

$$\max(w_n^0(a_1) - w_n^0(b_1), w_n^0(b_2) - w_n^0(a_2)) \geq \epsilon. \tag{1.8}$$

The effect of (1.8) on the eigenvalues of $(1/N)\mathbf{R}_n\mathbf{R}_n^*$ are immediate. Let $\lambda_{n1}, \leq \lambda_{n2}, \ldots$ denote the eigenvalues of $(1/N)\mathbf{R}_n\mathbf{R}_n^*$. Then, for all n large, there exists an index i_n for which

$$\lambda_{ni_n} < w_n^0(b_1) \quad \text{and} \quad \lambda_{n(i_n+1)} > w_n^0(b_2). \tag{1.9}$$

We impose another condition. Let $\mathbf{R}_{nj}$ denote the matrix resulting from removing the j-th column from $\mathbf{R}_n$. Then we require:

$$\begin{array}{l}\text{There exists a positive } \hat{\epsilon} \leq \epsilon \text{ and a positive } d < 1 \text{ such that for all } n \\ \text{large the number of } j\text{'s with no eigenvalues of } (1/(N-1))\mathbf{R}_{nj}\mathbf{R}_{nj}^* \\ \text{appearing in } w_n^0((a_1 - \hat{\epsilon}, a_2 + \hat{\epsilon})) \text{ is greater than } N^{1-d}.\end{array} \tag{1.10}$$

The main result can now be stated.

Theorem 1.1 *Assume*

(a) X_{ij}, $i,j = 1,2,\ldots$ *are independent standardized random variables.*

(b) *There exists a* K *and a random variable* X *with finite fourth moment such that, for any* $x > 0$

$$\frac{1}{n_1 n_2} \sum_{i \le n_1, j \le n_2} \mathrm{P}(|X_{ij}| > x) \le K\mathrm{P}(|X| > x)$$

for any n_1, n_2

(c) *There exists a positive function* $\psi(x) \uparrow \infty$ *as* $x \to \infty$*, and* $M > 0$ *such that*

$$\max_{ij} \mathrm{E}|X_{ij}^2|\psi(|X_{ij}|) \le M.$$

(d) *For* $n = 1, 2\ldots$ $\mathbf{R}_n$ *is an* $n \times N$ *nonrandom matrix with* $\|(1/\sqrt{N})\mathbf{R}_n\|$ *uniformly bounded for all* n.

(e) $N = N(n)$ *with* $c_n \equiv n/N \to c > 0$ *as* $n \to \infty$.

(f) $\mathbf{C}_n \equiv (1/N)(\mathbf{R}_n + \sigma\mathbf{X}_n)(\mathbf{R}_n + \sigma\mathbf{X}_n)^*$, *where* $\mathbf{X}_n = (X_{ij})$, $i = 1,2,\ldots,n$ $j = 1,2,\ldots,N$.

(g) *With* $b_1 > 0$, $[a_1, a_2] \subset (b_1, b_2)$ *with conditions (1.7) and (1.10) holding.*

Then $\mathrm{P}($*no eigenvalue of* $\mathbf{C}_n$ *appears in* $[a_1, a_2]$ *for all* n *large*$) = 1$.

Assumptions (a)-(c) allow for the extension of the X_{ij} to depart from merely being i.i.d. They were made in Couillet et. al. (2011), which extend the results in Bai and Silverstein (1998), (1999) on matrices of the form $(1/N)\mathbf{T}_n\mathbf{X}_n\mathbf{X}_n^*$, where $\mathbf{T}_n$ is Hermitian nonnegative definite, to non i.i.d. entries in $\mathbf{X}_n$. The argument in Couillet et. al. will be revisited in Section 3, which after suitable truncation, centralization, and scaling of the X_{ij}'s one can assume these variables to be uniformly bounded.

The rest of the paper will be devoted to proving the following.

Theorem 1.2 *Let* $z = x + iv_n$ *with* $v_n = \kappa n^{-1/(2p)}$*, where* p *is a positive integer. Then for* p *sufficiently large*

$$\sup_{x \in [a_1, a_2]} nv_n|m_n(z) - m_n^0| = o(1), \quad a.s.$$

As shown in Bai and Silverstein (1998), Theorem 1.2 proves Theorem 1.1. The main steps are given here. Let $[a_1', a_2'] \subset (b_1, b_2)$ for which $a_1' < a_1$, $a_2' > a_2$. Then from Theorem 1.2 it is straightforward to argue (with $I(A)$ denoting the indicator function on the set A)

$$\sup_{x \in [a_1, a_2]} \left| \int \frac{I([a_1', a_2']^c) d(F^{\mathbf{C}_n}(\lambda) - F^{c_n, H_n}(\lambda))}{((x-\lambda)^2 + v_n^2)((x-\lambda)^2 + 2v_n^2) \cdots ((x-\lambda)^2 + pv_n^2)} \right.$$
$$\left. + \sum_{\lambda_j \in [a_1', a_2']} \frac{v_n^{2p}}{((\lambda))((x-\lambda)^2 + v_n^2)((x-\lambda)^2 + 2v_n^2) \cdots ((x-\lambda)^2 + pv_n^2)} \right| = o(1), \quad a.s.,$$

where the λ_j's are the eigenvalues of $\mathbf{C}_n$. Since the integral converges a.s. to zero, one can argue, by contradiction that, a.s., that there can be no eigenvalues of $\mathbf{C}_n$ in $[a_1, a_2]$ for all n large.

The results in this paper are partway toward establishing the exact separation of the eigenvalues, that is, associating the eigenvalues, say, to the left of a_1 with eigenvalues of $(1/N)\mathbf{R}_n\mathbf{R}_n^*$ to the left of $w(a_1)$. Work is currently being pursued in this direction. It is remarked here that exact separation of eigenvalues of $\mathbf{C}_n$ has been achieved when the underlying distribution of the X_{ij} is complex normal, that is, real and imaginary parts are i.i.d. N(0,1/2) (Vallet, et. al.), using properties specific to this distribution. Work on exact separation of other ensembles of random matrices can be seen in Bai and Silverstein (1998), (1999), Capitaine, et.al. (2009) and Haagerup, et. al. (2006)

With the techniques used in this paper the authors found it necessary to include condition (1.10). It is needed to prove Lemma 3.2, involving the boundedness of moments of quadratic forms of resolvents of $\mathbf{C}_{(j)} = \mathbf{C}_n - (1/N)(\mathbf{r}_j + \sigma\mathbf{x}_j)(\mathbf{r}_j + \sigma\mathbf{x}_j)^*$, where $\mathbf{r}_j$ and $\mathbf{x}_j$ are the respective j-th columns of $\mathbf{R}_n$ and $\mathbf{X}_n$. This in turn required the absence of eigenalues of $(1/(N-1))\mathbf{R}_{nj}\mathbf{R}_{nj}^*$ in the interval (c_1, c_2). As an example, (1.10) disallows the inclusion of $\mathbf{R}_n$ constructed the following way. Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the canonical basis set in $\mathbb{R}^n$, and let $N \geq 2n$. Let, for $j \leq n$, $\mathbf{r}_j = \sqrt{Nj/n}\mathbf{e}_j$, $\mathbf{r}_{n+j} = \sqrt{N(1-j/n)}\mathbf{e}_{n+j}$, and $\mathbf{r}_j = \mathbf{0}$ for $j > 2n$. Then $(1/N)\mathbf{R}_n\mathbf{R}_n^* = \mathbf{I}$, the identity matrix, but for $j \leq n$ $(1/(N-1))\mathbf{R}_{nj}\mathbf{R}_{nj}^*$ has eigenvalue $1 - j/n$, which fill up $(0,1)$. Condition (1.10) is not needed when the X_{ij} are Gaussian, because in this case the distribution of $\mathbf{X}_n$ is invariant under unitary transformations, so it can be assumed that $\mathbf{R}_j$ is diagonal. Work is proceeding on removing this condition.

The proof of Theorem 1.2 will be done in Sections 3-5, where Sections 6,7 contain the proofs of lemmas need in the previous sections. Along the way will be derivations of bounds on moments of quadratic forms involving the resolvent of $\mathbf{C}_n$. Work in this area has been done in Hachem, et. al. As in our derivation, the bounds in that paper are in terms of powers of the reciprocal of the imaginary part of z. However the relation between moments and the powers of the reciprocal are unspecified, whereas it is important in our proof to know the dependence between the two. Therefore, it is necessary to include the derivations.

The authors are indebted to Philippe Loubaton, Walid Hachem, Malika Kharouf, Jamal Najim, and Pascal Vallet, for their help with crucial parts in the proof.

We begin by listing results needed in the proof of Theorem 1.2.

2 Mathematical tools.

Lemma 2.1 *(Lemma 3.3 of Dozier and Silverstein (2007a)). Let $z \in \mathbb{C}^+$ with $v = Im\, z$, $\mathbf{A}$ and $\mathbf{B}$ $n \times n$ with $\mathbf{B}$ Hermitian, and $\mathbf{r} \in \mathbb{C}^n$. Then*

$$|\mathrm{tr}((\mathbf{B} - z\mathbf{I})^{-1} - (\mathbf{B} + \mathbf{r}\mathbf{r}^* - z\mathbf{I})^{-1})\mathbf{A}| = \left| \frac{\mathbf{r}^*(\mathbf{B} - z\mathbf{I})^{-1}\mathbf{A}(\mathbf{B} - zI)^{-1}\mathbf{r}}{1 + \mathbf{r}^*(\mathbf{B} - z\mathbf{I})^{-1}r} \right| \leq \frac{\|\mathbf{A}\|}{v}.$$

Lemma 2.2 *For $\mathbf{a}, \mathbf{b} \in \mathbb{C}^n$, and $n \times n$ for which $\mathbf{A}$ and $\mathbf{A} + (\mathbf{a} + \mathbf{b})(\mathbf{a} + \mathbf{b})^*$ are both invertible we have*

$$\begin{aligned} &\mathbf{a}^*(\mathbf{A} + (\mathbf{a} + \mathbf{b})(\mathbf{a} + \mathbf{b})^*)^{-1} \\ &= \frac{1 + \mathbf{b}^*\mathbf{A}^{-1}(\mathbf{a} + \mathbf{b})}{1 + (\mathbf{a} + \mathbf{b})^*\mathbf{A}^{-1}(\mathbf{a} + \mathbf{b})}\mathbf{a}^*\mathbf{A}^{-1} - \frac{\mathbf{a}^*\mathbf{A}^{-1}(\mathbf{a} + \mathbf{b})}{1 + (\mathbf{a} + \mathbf{b})^*\mathbf{A}^{-1}(\mathbf{a} + \mathbf{b})}\mathbf{b}^*\mathbf{A}^{-1}. \end{aligned}$$

Lemma 2.3 *(Lemma B.26 of Bai and Silverstein (2009)). For $\mathbf{X} = (X_1, \ldots, X_n)^T$, with X_i's independent, standardized, and bounded, $\mathbf{C}$ an $n \times n$ matrix, we have for any $p \geq 2$*

$$\mathrm{E}|\mathbf{X}^*\mathbf{C}\mathbf{X} - \mathrm{tr}\mathbf{C}|^p \leq K_p(\mathrm{tr}\mathbf{C}\mathbf{C}^*)^{p/2}.$$

where K_p also depends on the bound on the X_i's.

Corollary 2.1 *For $\mathbf{X}$ as in Lemma 2.3 and $p \geq 2$ we have*

$$\mathrm{E}\|\mathbf{X}\|^{2p} \leq K_p n^p.$$

Moreover, for $\mathbf{y} \in \mathbb{C}^n$ *nonrandom and* $p \geq 2$

$$\mathrm{E}|\mathbf{X}^*\mathbf{y}|^p \leq K_p\|\mathbf{y}\|^p.$$

Proof. We have, using Lemma 2.3

$$\mathrm{E}\|\mathbf{X}\|^{2p} = \mathrm{E}(\mathbf{X}^*\mathbf{I}\mathbf{X})^p \leq 2^{p-1}(\mathrm{E}|\mathbf{X}^*\mathbf{I}\mathbf{X} - \mathrm{tr}\mathbf{I}|^p + (\mathrm{tr}\mathbf{I})^p) \leq K_p(n^{p/2} + n^p) \leq K_p n^p.$$

Again, using Lemma 2.3

$$\mathrm{E}|\mathbf{X}^*\mathbf{y}|^p \leq (\mathrm{E}|\mathbf{X}^*\mathbf{y}\mathbf{y}^*\mathbf{X}|^p)^{1/2} \leq 2^{(p-1)/2}(\mathrm{E}|\mathbf{X}^*\mathbf{y}\mathbf{y}^*\mathbf{X} - \mathrm{tr}\mathbf{y}\mathbf{y}^*|^p + (\mathrm{tr}\mathbf{y}\mathbf{y}^*)^p)^{1/2}$$
$$\leq 2^{(p-1)/2}(K_p((\mathrm{tr}(\mathbf{y}\mathbf{y}^*)^2)^{p/2} + \|\mathbf{y}\|^{2p})^{1/2} \leq K_p\|\mathbf{y}\|^p.$$

Lemma 2.4 *(Corollary 7.3.8 of Horn and Johnson (1985)). For* $r \times s$ *matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *with respective singular values* $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_q$, $\tau_1 \geq \tau_2 \geq \cdots \geq \tau_q$, *where* $q = \min(r, s)$ *we have*

$$|\sigma_k - \tau_k| \leq \|\mathbf{B} - \mathbf{A}\| \quad \text{for all } k = 1, 2, \ldots, q.$$

Lemma 2.5 *(Lemma 2.2 of Shohat and Tamarkin (1970)). If* f *is analytic on* $\mathbb{C}^+$, $f(z)$ *maps* $\mathbb{C}^+$ *into* $\mathbb{C}^+$, *and there is a* $\theta \in (0, \pi/2)$ *for which* $zf(x) \to c$, *finite, as* $z \to \infty$ *restricted to* $\{w \in \mathbb{C}^+ : \theta < \arg w < \pi - \theta\}$, *then* $c < 0$ *and* f *is the Stieltjes transform of a measure on* $\mathbb{R}$ *with total mass* $-c$.

Lemma 2.6 *(Couillet et. al. (2011)). When the entries of* X_n *are bounded the largest eigenvalue of* $\frac{1}{N}X_nX_n^*$, *denoted by* $\lambda_{\max}^{(1/N)X_nX_n^*}$, *satisfies*

$$\mathrm{P}(\lambda_{\max}^{(1/N)X_nX_n^*} > K) = o(n^{-t})$$

for any $K > (1 + \sqrt{c})^2$ *and any positive* t.

Lemma 2.7 *(Problem 8, p. 17 of Billingsley (1968)). Suppose* P_n, P *are probability measures on a separable metric space* S *for which* P_n *converges weakly to* P *as* $n \to \infty$. *Let* $\{f_\theta\}$, $\theta \in \Theta$ *be a family of real valued uniformly bounded functions on* S *and equicontinuous at each* $x \in S$, *that is, for each* x *and* $\epsilon > 0$ *there exists a* $\delta > 0$ *for which* $|x - y| < \delta$ *implies, for all* $\theta \in \Theta$, $|f_\theta(x) - f_\theta(y)| < \epsilon$. *Then*

$$\lim_{n\to\infty} \sup_{\theta\in\Theta} \left| \int f_\theta d\mathrm{P}_n - \int f_\theta d\mathrm{P} \right| = 0.$$

Lemma 2.8 *If for all $\epsilon > 0$, $\mathrm{P}(|X| > \epsilon)\epsilon^p \le K$ for some positive p, then, for any positive $q < p$*

$$\mathrm{E}|X|^q \le K^{q/p}\left(\frac{p}{p-q}\right).$$

Proof. For any > 0

$$\mathrm{E}|X|^q = \int_0^\infty \mathrm{P}(|X|^q > t)dt \le a + K\int_a^\infty t^{-p/q}dt = a + K\frac{q}{p-q}a^{1-p/q}.$$

By differentiating the last expression with respect to a and setting it to zero, we find its minimum occurs when $a = K^{q/p}$, its value giving us the desired upper bound.

Lemma 2.9 *(Burkholder (1973)). Let $\{X_k\}$ be a complex martingale difference sequence with respect to the increasing σ-fields $\{\mathcal{F}_n\}$. Then, for $p \ge 2$*

$$\mathrm{E}\left|\sum X_k\right|^p \le K_p\left(\mathrm{E}\left(\sum \mathrm{E}(|X_k|^2|\mathcal{F}_{k-1})\right)^{p/2} + \mathrm{E}\sum |X_k|^p\right).$$

Lemma 2.10 *(Burkholder (1973)). Let $\{X_k\}$ be as above. Then for any $p > 1$*

$$\mathrm{E}\left|\sum X_k\right|^p \le K_p\mathrm{E}\left(\sum |X_k|^2\right)^{p/2}.$$

Lemma 2.11 *For Hermitian $n \times n$ Hermitian matrices $\mathbf{A}$ and $\mathbf{B}$, we have for any positive even integer ℓ*

$$\mathrm{tr}(\mathbf{AB})^\ell \le \|\mathbf{A}\|^\ell \mathrm{tr}\mathbf{B}^\ell$$

(follows from Theorem 3.2 of Yang and Feng (2002)).

3 Convergence of $m_n - \mathrm{E}m_n$ and certain quadratic forms.

Constants appearing in inequalities are designated by K, sometimes subscripted. They are nonrandom and may differ from one appearance to the next.

We begin with simplifying assumptions. Suppose $n > N$. Then the largest $n - N$ eigenvalues of C_n are identical to the eigenvalues of $(1/N)(\mathbf{R}_n + \sigma\mathbf{X}_n)^*(\mathbf{R}_n + \sigma\mathbf{X}_n) = (n/N)(1/n)(\mathbf{R}_n^* + \sigma\mathbf{X}_n^*)(\mathbf{R}_n^* + \sigma\mathbf{X}_n^*)^* = c_n\underline{\mathbf{C}}_n$, where $\underline{\mathbf{C}}_n$ is defined exactly in the same manner as $\mathbf{C}_n$ with the roles of n and N reversed. Since we are dealing with eigenvalues on the positive reals, the truth of Theorem 1.1 for limiting $c < 1$ will imply the truth for $c > 1$. So we will henceforth assume $c_n \leq 1$ for all n.

Since the eigenvalues of $\mathbf{C}_n$ and $\sigma^{-2}\mathbf{C}_n$ differ by a scaling factor, we may assume $\sigma = 1$.

We proceed with simplifying the assumptions on the entries of $\mathbf{X}_n$. We follow along the truncation and centralization steps taken on the entries of $\mathbf{X}_n$ in Appendix B-B of Couillet et. al. (2011), which involves matrices of the form $(1/N)(\mathbf{X}_n\mathbf{X}_n^*)\mathbf{T}_n$, where $\mathbf{T}_n$ is nonnegative definite. The conditions (a)-(c) made in Theorem 1.1 are identical to those made on the entries of $\mathbf{X}_n$ in Couillet et. al. (2011). In 2.-4. of that paper Lemma 2.4 is used on $(1/\sqrt{N})\mathbf{T}_n^{1/2}\mathbf{X}_n'$ and $(1/\sqrt{N})\mathbf{T}_n^{1/2}\mathbf{X}_n''$, where $\mathbf{X}_n''$ results in either a truncation, a centralization, or a scaling of the entries of $\mathbf{X}_n'$. We get the same results when we apply Lemma 2.4 on $(1/\sqrt{N})(\mathbf{R}_n + \mathbf{X}_n')$ and $(1/\sqrt{N})(\mathbf{R}_n + \mathbf{X}_n'')$. Thus we conclude that we may assume as in Couillet et. al. (2011) the entries of $\mathbf{X}_n$ are uniformly bounded.

Let

$$\widehat{\mathbf{A}}_n = \frac{1}{z(1 + c_n\mathrm{E}m_n(z))}(1/N)\mathbf{R}_n\mathbf{R}_n^* - \mathrm{E}\mathbf{m}_n(z)\mathbf{I},$$

where

$$\mathbf{m}_n(z) = m_{F^{(1/N)(\mathbf{R}_n+\mathbf{X}_n)^*(\mathbf{R}_n+\mathbf{X}_n)}}(z).$$

We have

$$\mathbf{m}_n(z) = -\frac{1 - c_n}{z} + c_n m_n(z). \tag{3.1}$$

Also, let $\mathbf{D} = \mathbf{D}(z) = \mathbf{C}_n - z\mathbf{I}$, and $\mathbf{u} \in \mathbb{C}^n$, with $\|\mathbf{u}\| = 1$. In this section we shall verify bounds on moments of

$$m_n - \mathrm{E}m_n \tag{3.2}$$

and on

$$\mathbf{u}^*\mathbf{D}^{-1}\mathbf{u} - \mathrm{E}\mathbf{u}^*\mathbf{D}^{-1}\mathbf{u}. \tag{3.3}$$

The next section will establish bounds on

$$\mathrm{E}m_n - (1/(zn))\mathrm{tr}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1} \tag{3.4}$$

and

$$\mathrm{E}\mathbf{u}^*\mathbf{D}^{-1}\mathbf{u} - (1/z)\mathbf{u}^*(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{u}. \tag{3.5}$$

Two separate bounds will be shown simultaneously on (3.2) and (3.4). One bound, which we call the "a" bound, will hold for all $z = x + iv_n$ for x lying in a bounded interval, $[e, f]$, of $\mathbb{R}^+$. The second bound, called the "b" bound, will hold for all z with $x \in [a_1, a_2]$, but which rely on two lemmas, one depending on the truth of the "a" bound, the other on the bounds derived for (3.3) and (3.5). All four quantities will each be split into a sum of several terms, where "a" and "b" bounds will be derived for each term involving (3.2) and (3.4). Section 5 will finish the proof of Theorem 1.2, and Section 6 will contain the proof of the two lemmas, along with a corollary to one of the lemmas.

The first lemma is hereby given.

Lemma 3.1

(a) We have $F_n \xrightarrow{\mathcal{D}} F$ *almost surely.*

(b) For any bounded interval $[e, f]$ *with* $e > 0$ *and any* $r \geq 1$

$$\lim_{n\to\infty} \mathrm{E}(v_n^{-r} \sup_{x\in[e,f]} |m_n - m_n^0|^r) = 0.$$

We will essentially need the following corollary. First we introduce the following notation. Write $\mathbf{s}_j = N^{-1/2}(\mathbf{r}_j + \mathbf{x}_j)$, where $\mathbf{r}_j$ and $\mathbf{x}_j$ are the respective j-th columns of $\mathbf{R}_n$ and $\mathbf{X}_n$. Notice that the $\mathbf{s}_j$'s are uniformly bounded.

Corollary 3.1

(a) For positive integer m, *nonnegative integers* ν_k, μ_k, $k = 1, \ldots, m$ *with* $\nu_k + \mu_k \geq 1$, $\sum_{k=1}^m (\nu_k + \mu_k) \leq 16$, *and* $\ell \geq 1$ *we have uniformly for all* $x \in [a_1, a_2]$

$$\mathrm{E}\left| \frac{1}{N} \sum_{j=1}^N \prod_{k=1}^m \mathbf{s}_j^* \mathbf{D}^{-\nu_k} \mathbf{D}^{-\mu_k *} \mathbf{s}_j \right|^\ell \leq K_\ell.$$

(b) For positive integer $m \leq 8$ *we have uniformly for* $x \in [a_1, a_2]$

$$\frac{1}{N} \mathrm{E}\mathrm{tr}\mathbf{D}^{-m}\mathbf{D}^{-m*} \leq K$$

Before stating the second lemma we introduce some more notation. Write $\mathbf{C}_{(j)} = \mathbf{C} - \mathbf{s}_j\mathbf{s}_j^*$, $\mathbf{D}_j = \mathbf{C}_{(j)} - z\mathbf{I} = \mathbf{D} - \mathbf{s}_j\mathbf{s}_j^*$, and

$$\beta_j = \beta_j(z) = \frac{1}{1 + \mathbf{s}_j^*\mathbf{D}_j^{-1}(z)\mathbf{s}_j}.$$

The second lemma can now be stated.

Lemma 3.2 *We have for any $\ell \geq 1$ and $x \in [a_1, a_2]$*

$$\mathrm{E}\left(\frac{1}{N}\sum_{j=1}^{N}|\beta_j|^{-4}\right)^{\ell} \leq K_\ell.$$

In this section we will show for $v_n = \kappa n^{-1/p}$ for p suitable large, and any $\ell \geq 1$

$$\mathrm{E}|m_n(z) - \mathrm{E}m_n(z)|^{2\ell} \leq \begin{cases} K_\ell v_n^{-8\ell} n^{-2\ell} & (3.6\mathrm{a}) \\ K_\ell n^{-2\ell}, & (3.6\mathrm{b}) \end{cases}$$

and

$$\mathrm{E}|\mathbf{u}^*(\mathbf{D}^{-1} - \mathrm{E}\mathbf{D}^{-1})\mathbf{u}|^{2\ell} \leq K v_n^{-10\ell} n^{-\ell/3}. \tag{3.7}$$

The "a" bound and the quadratic form bound holds uniformly for all $x \in [e, f]$. The "b" bound holds uniformly for all $x \in [a_1, a_2]$. Notice these bounds hold for all $\ell \geq 1$ once they are shown to be true for sufficiently large ℓ.

Before proceeding we introduce some more notation. Let

$$\hat\beta_j = \frac{1}{1 + N^{-1}(\mathrm{tr}\mathbf{D}_j^{-1} + \mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{r}_j)},$$

$$\alpha_j = \mathbf{s}_j^*\mathbf{D}_j^{-2}\mathbf{s}_j - (1/N)(\mathrm{tr}\mathbf{D}_j^{-2} + \mathbf{r}_j^*\mathbf{D}_j^{-2}\mathbf{r}_j), \quad \eta_j = N^{-1}(\mathrm{tr}\mathbf{D}_j^{-2} + \mathbf{r}_j^*\mathbf{D}_j^{-2}\mathbf{r}_j),$$

and

$$\gamma_j = \mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{s}_j - N^{-1}(\mathrm{tr}\mathbf{D}_j^{-1} + \mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{r}_j).$$

From Lemma 2.5 we see that both $-z^{-1}\beta_j$ and $-z^{-1}\hat\beta_j$ are Stieltjes transforms of (random) probability measures. Consequently, β_j and $\hat\beta_j$ are bounded in absolute value by $|z|v_n^{-1}$. Notice, because $\|(1/N)\mathbf{R}_n\mathbf{R}_n^*\|$ and the entries of $\mathbf{X}_n$ are uniformly bounded, the vectors $\mathbf{s}_j$ and $N^{-1/2}\mathbf{r}_j$ are uniformly bounded in Eudlidean norm. Thus, for all n, $j \le N$, and $x \in [e, f]$

$$\max(|\beta_j|, |\hat{\beta}_j|, |\gamma_j|) \le K v_n^{-1}$$

and

$$\max(|\alpha_j|, |\eta_j|) \le K v_n^{-2}.$$

Since $\hat{\beta}_j - \beta_j = \gamma_j \beta_j \hat{\beta}_j$ we have

$$|\gamma_j \beta_j \hat{\beta}_j| \le K v_n^{-1}.$$

We also have

$$|\eta_j \hat{\beta}_j| \le v_n^{-1}. \tag{3.8}$$

Indeed, denote the spectral decomposition of $\mathbf{C}_{(j)}$ by $\mathbf{U}\mathbf{\Lambda}\mathbf{U}^*$, the eigenvalues, λ_{jk}, in $\mathbf{\Lambda}$, arranged in increasing order. Let $\mathbf{y} = \mathbf{U}^*\mathbf{r}_j = (y_1, \ldots, y_n)^T$. Then

$$|\eta_j| \le N^{-1} \sum_{k=1}^{n} \frac{1 + |y_k|^2}{|\lambda_{jk} - z|^2},$$

whereas

$$|1 + N^{-1}(\mathrm{tr}\mathbf{D}_j^{-1} + \mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{r}_j| \ge \Im N^{-1}\mathrm{tr}\mathbf{D}_j^{-1} + \mathbf{r}_j^*\mathbf{D}_j^{-1} r_j = N^{-1} v_n \sum_{k=1}^{n} \frac{1 + |y_k|^2}{|\lambda_{jk} - z|^2}.$$

Thus, (3.8) follows.

We first establish an essential bound. Let $\mathrm{E}_{(j)}$ denote conditional expectation with respect to $\mathbf{x}_1, \ldots, \mathbf{x}_{j-1}, \mathbf{x}_{j+1}, \ldots, \mathbf{x}_N$. Using Lemma 2.3 and Corollary 2.1, we have for $\ell \ge 1$

$$\begin{aligned}
\mathrm{E}_{(j)}|\gamma_j|^{2\ell} &\le K_\ell N^{-2\ell}(\mathrm{E}_{(j)}|\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{x}_j - \mathrm{tr}\mathbf{D}_j^{-1}|^{2\ell} + \mathrm{E}_{(j)}|\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{r}_j|^{2\ell} + \mathrm{E}_{(j)}|\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{x}_j|^{2\ell}) \\
&\le K_\ell N^{-2\ell}(\mathrm{E}_{(j)}(\mathrm{tr}\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*})^\ell + \mathrm{E}_{(j)}\|\mathbf{D}_j^{-1}\mathbf{r}_j\|^{2\ell} + \mathrm{E}_{(j)}\|\overline{\mathbf{D}}_j^{-1}\mathbf{r}_j\|^{2\ell}) \\
&\le K_\ell N^{-2\ell} N^\ell v_n^{-2\ell} \le K_\ell n^{-\ell} v_n^{-2\ell}.
\end{aligned} \tag{3.9}$$

Similarly we get

$$\mathrm{E}_{(j)}|\alpha_j|^{2\ell} \le K_\ell n^{-\ell} v_n^{-4\ell}. \tag{3.10}$$

Restricting $v_n = \kappa n^{-1/p}$ with $p > 10$ we have for $\ell \ge 2$

$$\mathrm{E}_{(j)}|\gamma_j|^\ell \le K_\ell n^{-2\ell/5}. \tag{3.11}$$

Let $I(A)$ denote the indicator on the event A. Then from (3.11) we have for any $t > 0$

$$\mathrm{P}_{(j)}(|\gamma_j| > n^{-1/3}) = \mathrm{E}_{(j)} I(|\gamma_j| > n^{-1/3}) = o(n^{-t}). \tag{3.12}$$

Here, $\mathrm{P}_{(j)}$ is the conditional probability with respect to $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_N$. These bounds are uniform across j. We have then for any $t > 0$

$$\mathrm{P}(\sup_{j \le N} |\gamma_j| > n^{-1/3}) = o(n^{-t}). \tag{3.13}$$

We will need the following.

Lemma 3.3 *Suppose f is a random variable for which $|f| \le n^{\mu}$ for some positive μ. Then for any $j \le N$, $\eta \ge 1$ and $t > 0$*

$$\mathrm{E}_{(j)} |\beta_j - \hat{\beta}_j|^{\eta} |f| \le K n^{-\eta/3} v_n^{-2\eta} \mathrm{E}_{(j)} |f| + o(n^{-t}).$$

The bound K depends only on η and the interval $[e, f]$.

Proof. We have by (3.12)

$$\mathrm{E}_{(j)} |\beta_j - \hat{\beta}_j|^{\eta} |f| \le K v_n^{-\eta} n^{\mu} \mathrm{E}_{(j)} I(|\gamma_j| > n^{-1/3}) + K n^{-\eta/3} v_n^{-2\eta} \mathrm{E}_{(j)} |f|$$

$$\le K v_n^{-\eta} n^{\mu} \mathrm{P}_{(j)}(|\gamma_j| > n^{-1/3}) + K n^{-\eta/3} v_n^{-2\eta} \mathrm{E}_{(j)} |f| \le K n^{-\eta/3} v_n^{-2\eta} \mathrm{E}_{(j)} |f| + o(n^{-t}).$$

As a consequence of this lemma, we prove the following useful bound:

$$\max\left(\mathrm{E}\Big(\sum_{j=1}^{N} \mathrm{E}_{j-1} |\beta_j \gamma_j|^2 \Big)^{\ell}, \mathrm{E}\Big(\sum_{j=1}^{N} \mathrm{E}_{j-1} |\hat{\beta}_j \gamma_j|^2 \Big)^{\ell} \right) \le \begin{cases} K_\ell v_n^{-2\ell} & \text{(3.14a)} \\ K_\ell. & \text{(3.14b)} \end{cases}$$

Indeed, using Lemma 3.3 and the identity $\mathbf{s}_j^* \mathbf{D}^{-1} = \beta_j \mathbf{s}_j \mathbf{D}_j^{-1}$ we have

$$\mathrm{E}\left(\sum_{j=1}^{N} \mathrm{E}_{j-1} |\hat{\beta}_j \gamma_j|^2 \right)^{\ell} = \mathrm{E}\left(\frac{1}{N^2} \sum_{j=1}^{N} \mathrm{E}_{j-1} |\hat{\beta}_j|^2 (\mathbf{r}_j^* \mathbf{D}_j^{-1} \mathbf{D}_j^{-1*} \mathbf{r}_j + \mathrm{tr}(\mathbf{D}_j^{-1} \mathbf{D}_j^{-1*})) \right)^{\ell}$$

$$= \mathrm{E}\left(\frac{1}{N} \sum_{j=1}^{N} \mathrm{E}_{j-1} |\hat{\beta}_j|^2 \mathbf{s}_j^* \mathbf{D}_j^{-1} \mathbf{D}_j^{-1*} \mathbf{s}_j \right)^{\ell}$$

$$\leq K_\ell \mathrm{E}\left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}_{j-1}|\beta_j|^2\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{s}_j\right)^\ell + K_\ell \mathrm{E}\left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}_{j-1}|\hat\beta_j-\beta_j|^2\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{s}_j\right)^\ell$$

$$\leq K_\ell \mathrm{E}\left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}_{j-1}\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{D}^{-1*}\mathbf{s}_j\right)^\ell + (n^{-2/3}v_n^{-6})^\ell$$

$$\leq K_\ell \mathrm{E}\left(\frac{1}{N}\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{D}^{-1*}\mathbf{s}_j\right)^\ell + o(1)$$

from which we immediately get (3.14a) and, from Corollary 3.1, (3.14b) for the second quantity on the left in (3.14). For the first quantity we have

$$\mathrm{E}\Big(\sum_{j=1}^N \mathrm{E}_{j-1}|\beta_j\gamma_j|^2\Big)^\ell \leq K_\ell\mathrm{E}\Big(\sum_{j=1}^N \mathrm{E}_{j-1}|\hat\beta_j\gamma_j|^2\Big)^\ell + K_\ell\mathrm{E}\Big(\sum_{j=1}^N \mathrm{E}_{j-1}|\beta_j-\hat\beta_j|^2|\gamma_j|^2\Big)^\ell,$$

where from Lemma 3.3

$$\mathrm{E}\Big(\sum_{j=1}^N \mathrm{E}_{j-1}|\beta_j-\hat\beta_j|^2|\gamma_j|^2\Big)^\ell \leq K_\ell n^{-2\ell/3}v_n^{-4\ell}\mathrm{E}\Big(\sum_{j=1}^N \mathrm{E}_{j-1}|\gamma_j|^2\Big)^\ell + o(n^{-t})$$

$$= K_\ell n^{-2\ell/3}v_n^{-4\ell}\mathrm{E}\Big(\frac{1}{N^2}\sum_{j=1}^N \mathrm{E}_{j-1}(\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{r}_j + \mathrm{tr}(\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}))\Big)^\ell + o(n^{-t})$$

$$\leq K_\ell n^{-2\ell/3}v_n^{-6\ell} \to 0$$

as $n \to \infty$. Therefore we get the bound on the first quantity on the left in (3.14).

We proceed now with (3.6). Let $\mathrm{E}_0(\cdot)$ denote expectation and $\mathrm{E}_k(\cdot)$ denote conditional expectation with respect to the σ-field generated by the first k columns of the doubly infinite matrix from which the $\mathbf{X}_n$'s are taken from. Using Lemma 2.1 we write

$$n(\mathrm{E}m_n(z) - m_n(z)) = -\sum_{j=1}^N(\mathrm{E}_j\mathrm{tr}\mathbf{D}^{-1} - \mathrm{E}_{j-1}\mathbf{D}^{-1})$$

$$= -\sum_{j=1}^N(\mathrm{E}_j(\mathrm{tr}\mathbf{D}^{-1} - \mathrm{tr}\mathbf{D}_j^{-1}) - \mathrm{E}_{j-1}(\mathrm{tr}\mathbf{D}^{-1} - \mathrm{tr}\mathbf{D}_j^{-1}))$$

$$= \sum_{j=1}^N(\mathrm{E}_j - \mathrm{E}_{j-1})\mathbf{s}_j^*\mathbf{D}_j^{-2}\mathbf{s}_j\beta_j.$$

Expand β_j into

$$\beta_j = \hat{\beta}_j - \gamma_j \hat{\beta}_j^2 + \gamma_j^2 \hat{\beta}_j^2 \beta_j. \tag{3.15}$$

We have then

$$n(\mathrm{E}m_n(z) - m_n(z)) = J_1 - J_2 - J_3 + J_4,$$

where

$$\begin{aligned}
J_1 &= \sum_{j=1}^{N} (\mathrm{E}_j - \mathrm{E}_{j-1}) \alpha_j \hat{\beta}_j, \\
J_2 &= \sum_{j=1}^{N} (\mathrm{E}_j - \mathrm{E}_{j-1}) \alpha_j \hat{\gamma}_j \hat{\beta}_j^2 \\
J_3 &= \sum_{j=1}^{N} (\mathrm{E}_j - \mathrm{E}_{j-1}) \eta_j \hat{\gamma}_j \hat{\beta}_j^2 \\
J_4 &= \sum_{j=1}^{N} (\mathrm{E}_j - \mathrm{E}_{j-1}) \mathbf{s}_j^* \mathbf{D}_j^{-2} \mathbf{s}_j \hat{\gamma}_j^2 \hat{\beta}_j^2 \beta_j
\end{aligned}$$

Note that each J_k is a sum of martingale differences. Our goal is to show for $\ell \geq 1$

$$\mathrm{E}|J_k|^{2\ell} \leq \begin{cases} K_\ell v_n^{-8\ell} & (3.16\mathrm{a}) \\ K_\ell, & (3.16\mathrm{b}) \end{cases}$$

for $k = 1, 2, 3, 4$.

We begin with J_2 and J_4. Write, for $k = 1, 2$, $J_k = J_{k1} + J_{k2}$, where J_{k1} (J_{k2}) results in placing $I(|\gamma_k| < N^{-1/3})$ ($I(|\gamma_k| \geq N^{-1/3})$) to the right of $(\mathrm{E}_j - \mathrm{E}_{j-1})$ in each term of J_k. From (3.13) we have for $\ell \geq 1$ and all positive t

$$\begin{aligned}
\mathrm{E}\,|J_{k2}|^{2\ell} &\leq K_\ell \mathrm{E} \left(\sum_{j=1}^{N} v_n^{-5} I(|\gamma_j| \geq N^{-1/3}) \right)^{2\ell} \\
&\leq K_\ell N^{2\ell-1} v^{-10\ell} \sum_{j=1}^{N} \mathrm{P}(|\gamma_j| \geq N^{-1/3}) \leq K_\ell N^{2l-t} v_n^{-10\ell} \to 0. \qquad (3.17)
\end{aligned}$$

as $n \to \infty$. All subsequent moment bounds will use Lemma 2.9. For J_{41} we have using (3.14a)

$$\mathrm{E}|J_{41}|^{2\ell} \leq K_\ell \Bigg[\mathrm{E} \Bigg(\sum_{j=1}^{N} \mathrm{E}_{j-1} \left| (\mathrm{E}_j - \mathrm{E}_{j-1}) \mathbf{s}_j^* \mathbf{D}_j^{-2} \mathbf{s}_j \gamma_j^2 \hat{\beta}_j^2 \beta_j I(|\gamma_j| < n^{-1/3}) \right|^2 \Bigg)^\ell$$

$$+\sum_{j=1}^{N}\mathrm{E}\left|(\mathrm{E}_j-\mathrm{E}_{j-1})\mathbf{s}_j^*\mathbf{D}_j^{-2}\mathbf{s}_j\gamma_j^2\hat{\beta}_j^2\beta_jI(|\gamma_j|<n^{-1/3})\right|^{2\ell}\Bigg]$$

$$\leq K_\ell\Big[v_n^{-8\ell}n^{-2\ell/3}\mathrm{E}\Big(\sum_{j=1}^{N}\mathrm{E}_{j-1}|\hat{\beta}_j\gamma_j|^2\Big)^\ell+v_n^{-10\ell}n^{-4\ell/3+1}\Big]$$

$$\leq K_\ell v_n^{-10\ell}(n^{-2\ell/3}+n^{-4\ell/3+1})\to 0$$

(when $p>30$) as $n\to\infty$.

We have

$$\mathrm{E}|J_{21}|^{2l}\leq K_\ell n^{-2\ell/3}v_n^{-4\ell}\mathrm{E}\left(\sum_{j=1}^{N}\mathrm{E}_{j-1}|\alpha_j|^2\right)^\ell+K_\ell N^{-2\ell/3+1}v_n^{-8\ell}$$

$$\leq \quad K_\ell n^{-2\ell/3}v_n^{-4\ell}\mathrm{E}\left(N^{-2}\sum_{j=1}^{N}\mathrm{E}_{j-1}\Big(\mathrm{tr}\mathbf{D}_j^{-2}\mathbf{D}_j^{-2*}+\mathbf{r}_j^*\mathbf{D}_j^{-2}\mathbf{D}_j^{-2*}\mathbf{r}_j\Big)\right)^l+K_\ell n^{-2l/3+1}v^{-8\ell}$$

$$\leq \quad K_\ell(n^{-2\ell/3}v_n^{-8\ell}+n^{-2l/3+1}v_n^{-8\ell})\leq K_\ell n^{-2\ell/3+1}v_n^{-8\ell}\to 0.,$$

for $\ell\geq 2$ and $p>48$. Thus we have (3.14) for $k=4$ and 2.

We next handle J_3. Using (3.11), (3.8), Lemma 3.3, and

$$|\mathrm{tr}\mathbf{D}_j^{-2}+\mathbf{r}_j^*\mathbf{D}_j^{-2}\mathbf{r}_j|\leq \mathrm{tr}\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}+\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{r}_j,$$

we have for $\ell\geq 2$

$$\mathrm{E}|J_3|^{2\ell}\leq K_\ell\left[\mathrm{E}\left(\sum_{j=1}^{N}\mathrm{E}_{j-1}|\eta_j\gamma_j|^2|\hat{\beta}_j|^4\right)^\ell+N^{-4\ell/5+1}v_n^{-4\ell}\right]$$

$$\leq K_\ell\left[\mathrm{E}\left(\frac{1}{N^2}\sum_{j=1}^{N}\mathrm{E}_{j-1}\Big(\mathrm{tr}\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}+\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{r}_j\Big)|\eta_j^2\hat{\beta}_j^4|\right)^\ell+N^{-4\ell/5+1}v_n^{-4\ell}\right]$$

$$\leq K_\ell\left[\mathrm{E}\left(\frac{1}{N^4}\sum_{j=1}^{N}\mathrm{E}_{j-1}\Big(\mathrm{tr}\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}+\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{r}_j\Big)^3|\hat{\beta}_j^4|\right)^\ell N^{-4\ell/5+1}v_n^{-4\ell}\right]$$

$$=K_\ell\mathrm{E}\left(\frac{1}{N}\sum_{j=1}^{N}\mathrm{E}_{j-1}\Big(\mathrm{E}_{(j)}\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{s}_j\Big)^3|\hat{\beta}_j|^4\right)^\ell+o(1)$$

$$\le K_\ell \mathrm{E}\Big(\frac{1}{N}\sum_{j=1}^N \mathrm{E}_{j-1}\Big(\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{s}_j\Big)^3|\beta_j|^4 + n^{-4/3}v_n^{-14}\Big)^\ell + o(1)$$

$$\le K_\ell \mathrm{E}\Big(\frac{1}{N}\sum_{j=1}^N \mathrm{E}_{j-1}\Big(\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{s}_j\Big)^3|\beta_j|^4\Big)^\ell + o(1).$$

$$= K_\ell \mathrm{E}\left(\frac{1}{N}\sum_{j=1}^N (\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{D}^{-1*}\mathbf{s}_j)^3|\beta_j^{-2}|\right)^\ell + o(1)$$

$$\le K_\ell\left[\mathrm{E}\left(\frac{1}{N}\sum_{j=1}^N (\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{D}^{-1*}\mathbf{s}_j)^6\right)^\ell\right]^{1/2}\left[\mathrm{E}\left(\frac{1}{N}\sum_{j=1}^N |\beta_j|^{-4}\right)^\ell\right]^{1/2} \qquad (3.18)$$
$$+o(1).$$

Since $|\beta_j^{-1}| = |1+\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{s}_j| \le Kv_n^{-1}$ we see that (3.16a) holds, and from Corollary 3.1 and Lemma 3.2 we see that (3.16b) is true.

Lastly, we handle J_1. Using (3.10) we have for $\ell \ge 2$

$$\mathrm{E}|J_1|^{2\ell} \le K_\ell \mathrm{E}\left(\frac{1}{N^2}\sum_{j=1}^N (\mathrm{E}_{j-1}\mathrm{tr}(\mathbf{D}_j^{-2}\mathbf{D}_j^{-2*}) + \mathbf{r}_j^*\mathbf{D}_j^{-2}\mathbf{D}_j^{-2*}\mathbf{r}_j)|\hat\beta_j|^2\right)^\ell + K_\ell v_n^{-2\ell}\sum_{j=1}^N \mathrm{E}|\alpha_j|^{2\ell}$$

$$\le K_\ell \mathrm{E}\left(\frac{1}{N}\sum_{j=1}^N (s_j^* D_j^{-2}D_j^{-2*}s_j)|\hat\beta_j|^2\right)^\ell + K_\ell v_n^{-6\ell}N^{-\ell+1} = K_\ell J_{11} + o(1).$$

We see that (3.16a) is satisfied. For the "b" bound, taking a similar approach used above for J_3, we obtain

$$J_{11} \le \mathrm{E}\left(\frac{1}{N}\sum_{j=1}^N (\mathbf{s}_j^*\mathbf{D}_j^{-2}\mathbf{D}_j^{-2*}\mathbf{s}_j)|\beta_j|^2\right)^\ell + o(1)$$

$$= \mathrm{E}\left(\frac{1}{N}\sum_{j=1}^N \mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{D}^{-1*}\mathbf{s}_j\right)^\ell + o(1)$$

Using the formula $\mathbf{D}_j^{-1} = \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-1}\beta_j^{-1}$, and the inequality

$$(\mathbf{a}+\mathbf{b})^*(\mathbf{a}+\mathbf{b}) \le 2\mathbf{a}^*\mathbf{a} + 2\mathbf{b}^*\mathbf{b}$$

with $\mathbf{a} = \mathbf{D}^{-2*}\mathbf{s}_j$, $\mathbf{b} = \bar{\beta}_j^{-1}\mathbf{D}^{-1*}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-2*}\mathbf{s}_j$, we obtain

$$J_{11} \le K_\ell \mathrm{E}\left(\frac{1}{N}\sum_{j=1}^N \mathbf{s}_j^*\mathbf{D}^{-2}\mathbf{D}^{-2*}\mathbf{s}_j\right)^\ell + K_\ell \mathrm{E}\left(\frac{1}{N}\sum_{j=1}^N \mathbf{s}_j^*\mathbf{D}^{-2}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{D}^{-1*}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-2*}\mathbf{s}_j|\beta_j|^{-2}\right)^\ell$$

$$\le K_\ell \mathrm{E}\left(\frac{1}{N} s_j^*\mathbf{D}^{-2}\mathbf{D}^{-2*}\mathbf{s}_j\right)^\ell$$

$$+K_\ell\left[\mathrm{E}\left(\frac{1}{N}\sum_{j=1}^N(\mathbf{s}_j^*\mathbf{D}^{-2}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{D}^{-1*}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-2*}\mathbf{s}_j)^2\right)^\ell\right]^{1/2}\left[\mathrm{E}\left(\frac{1}{N}\sum_{j=1}^N|\beta_j|^{-4}\right)^\ell\right]^{1/2}.$$

So, from Corollary 3.1 and Lemma 3.2 we have we have (3.16b) for J_1, so that (3.6) holds.

We proceed to verify (3.7). Using (3.15) we expand in terms of sums of martingale differences:

$$\mathbf{u}^*(\mathrm{E}\mathbf{D}^{-1} - \mathbf{D}^{-1})\mathbf{u} = -\sum_{j=1}^N(\mathrm{E}_j - \mathrm{E}_{j-1})\mathbf{u}^*(\mathbf{D}^{-1} - \mathbf{D}_j^{-1})\mathbf{u}$$

$$= \sum_{j=1}^N(\mathrm{E}_j - \mathrm{E}_{j-1})\mathbf{u}_j^*\mathbf{D}_j^{-1}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{u}_j\beta_j = \breve{J}_1 - \breve{J}_2 - \breve{J}_3 + \breve{J}_4,$$

where

$$\begin{aligned}
\breve{J}_1 &= \sum_{j=1}^N(\mathrm{E}_j - \mathrm{E}_{j-1})\breve{\alpha}_j\hat{\beta}_j,\\
\breve{J}_2 &= \sum_{j=1}^N(\mathrm{E}_j - \mathrm{E}_{j-1})\breve{\alpha}_j\gamma_j\hat{\beta}_j^2\\
\breve{J}_3 &= \sum_{j=1}^N(\mathrm{E}_j - \mathrm{E}_{j-1})\breve{\eta}_j\gamma_j\hat{\beta}_j^2\\
\breve{J}_4 &= \sum_{j=1}^N(\mathrm{E}_j - \mathrm{E}_{j-1})\mathbf{u}_j^*\mathbf{D}_j^{-1}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{u}_j\gamma_j^2\hat{\beta}_j^2\beta_j.
\end{aligned}$$

Here,

$$\begin{aligned}
\breve{\alpha}_j &= \mathbf{u}^*\mathbf{D}_j^{-1}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{u} - \frac{1}{N}(\mathbf{u}^*\mathbf{D}_j^{-2}\mathbf{u} + \mathbf{u}^*\mathbf{D}_j^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}),\\
\breve{\eta}_j &= \frac{1}{N}(\mathbf{u}^*\mathbf{D}_j^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u} + \mathbf{u}^*\mathbf{D}_j^{-2}\mathbf{u}) = \mathrm{E}_{(j)}\mathbf{u}^*\mathbf{D}_j^{-1}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{u}.
\end{aligned}$$

Notice $\max(\breve{\alpha}_j, \breve{\eta}_j) \le K v_n^{-2}$, and from Lemma 2.3 and Corollary 2.1, for $\ell \ge 1$

$$\mathrm{E}_{(j)}|\breve{\alpha}_j|^{2\ell} \le N^{-2\ell}(|\mathbf{u}^*\mathbf{D}_j^{-1}\mathbf{r}_j|^{2\ell}(\mathbf{u}^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{u})^{\ell} + (\mathbf{u}^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{u})^{2\ell}) \le N^{-\ell}v_n^{-4\ell}. \quad (3.19)$$

As was done for J_2 and J_4 we split $\breve{J}_2$ and $\breve{J}_4$ according to whether $|\gamma_j|$ is less than or great than $n^{-1/3}$. Similar to the estimation on J_{k2}, we have, using (3.13), for any $t > 0$ and $\ell \ge 1$

$$\mathrm{E}|\breve{J}_{k2}|^{2\ell} \le K_\ell N^{2\ell - t} v^{-14\ell} \le K N^{-\ell} v_n^{-14\ell}, \quad k = 2, 4, \quad (3.20)$$

for t sufficiently large. By Lemma 2.10,

$$\begin{aligned} \mathrm{E}|\breve{J}_{41}|^{2\ell} &\le K_\ell \mathrm{E}\left(\sum_{j=1}^{N} |\mathbf{u}^*\mathbf{D}_j^{-1}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{u}\gamma_j^2\hat{\beta}_j^2\beta_j|^2 I(|\gamma_j| < n^{-1/3})\right)^l \\ &\le K N^{-\ell/3} v_n^{-10\ell}. \end{aligned} \quad (3.21)$$

The remaining bounds will use Lemma reflem28. We have from (3.19)

$$\begin{aligned} \mathrm{E}|\breve{J}_{21}|^{2\ell} &\le K_\ell \mathrm{E}\left(n^{-2/3}\sum_{j=1}^{N} \mathrm{E}_{j-1}|\breve{\alpha}_j|^2|\hat{\beta}_j|^4\right)^{\ell} + K_\ell n^{-2\ell/3} v_n^{-4\ell}\sum_{j=1}^{N}\mathrm{E}|\breve{\alpha}_j|^{2\ell} \\ &\le K_\ell N^{-2\ell/3} v_n^{-8\ell}. \end{aligned} \quad (3.22)$$

For $\breve{J}_3$ we have, using Lemma 3.3 and (3.11)

$$\begin{aligned} \mathrm{E}|\breve{J}_3|^{2\ell} &\le K_\ell \mathrm{E}\left(N^{-1}v_n^{-2}\sum_{j=1}^{N}\mathrm{E}_{j-1}|\breve{\eta}_j|^2|\hat{\beta}_j|^4\right)^{\ell} + K_\ell \sum_{j=1}^{N}\mathrm{E}|\breve{\eta}_j|^{2\ell}|\hat{\beta}_j|^{4\ell}|\gamma_j|^{2\ell} \\ &\le K_\ell N^{-\ell} v_n^{-2\ell}\left(\sum_{j=1}^{N}|\mathrm{E}_{(j)}(\mathbf{u}^*\mathbf{D}_j^{-1}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}_j^{-1*}\mathbf{u})|^2|\hat{\beta}_j^4|\right)^{\ell} + K_\ell N^{-4\ell/5+1} v_n^{-8\ell} \\ &\le K_\ell N^{-\ell} v_n^{-2\ell}\mathrm{E}\left(\sum_{j=1}^{N}|(\mathbf{u}^*\mathbf{D}_j^{-1}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}_j^{-1*}\mathbf{u})|^2|\beta_j|^4 + n^{-4/3}v_n^{-14}\right)^{\ell} + K_\ell N^{-4\ell/5+1} v_n^{-8\ell} \\ &\le K_\ell N^{-\ell} v_n^{-2\ell}\mathrm{E}\left(\sum_{j=1}^{N}|(\mathbf{u}^*\mathbf{D}^{-1}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-1*}\mathbf{u})|^2\right)^{\ell} + K_\ell N^{-4\ell/5+1} v_n^{-8\ell} \\ &\le K_\ell N^{-\ell} v_n^{-4\ell}\mathrm{E}\left(\sum_{j=1}^{N}(\mathbf{u}^*\mathbf{D}^{-1}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-1*}\mathbf{u})\right)^{\ell} + K_\ell N^{-4\ell/5+1} v_n^{-8\ell} \end{aligned}$$

$$
\begin{aligned}
&= K_\ell N^{-\ell} v^{-4\ell} \mathrm{E} \left(\mathbf{u}^* \mathbf{D}^{-1} \mathbf{C}_n \mathbf{D}^{-1*} \mathbf{u} \right)^\ell + K_\ell N^{-4\ell/5+1} v_n^{-8\ell} \\
&\le K_\ell N^{-\ell/3} v^{-8l},
\end{aligned}
\tag{3.23}
$$

for $\ell \ge 3$.

Using Lemma 3.3 and (3.19), we find

$$
\begin{aligned}
&\mathrm{E}|\breve{J}_1|^{2\ell} \le K_\ell \mathrm{E} \left(\sum_{j=1}^N \mathrm{E}_{j-1} |\breve{\alpha}_j|^2 |\hat{\beta}_j|^2 \right)^\ell + K_\ell \sum_{j=1}^N \mathrm{E} |\breve{\alpha}_j|^{2\ell} |\hat{\beta}_j|^{2\ell} \\
\le\ & K_\ell N^{-2\ell} \mathrm{E} \left(\sum_{j=1}^N (|\mathbf{u}^* \mathbf{D}_j^{-1} \mathbf{r}_j|^2 \mathbf{u}^* \mathbf{D}_j^{-1} \mathbf{D}_j^{-1*} \mathbf{u} + (\mathbf{u}^* \mathbf{D}_j^{-1} \mathbf{D}_j^{-1*} \mathbf{u})^2) |\hat{\beta}_j|^2 \right)^\ell \\
& + K_\ell N^{-2\ell} v_n^{-2\ell} \sum_{j=1}^N \mathrm{E} (|\mathbf{u}^* \mathbf{D}_j^{-1} \mathbf{r}_j|^{2\ell} (\mathbf{u}^* \mathbf{D}_j^{-1} \mathbf{D}_j^{-1*} \mathbf{u})^\ell + (\mathbf{u}^* \mathbf{D}_j^{-1} \mathbf{D}_j^{-1*} \mathbf{u})^{2\ell} \\
\le\ & K_\ell N^{-2\ell} v_n^{-2\ell} \mathrm{E} \left(\sum_{j=1}^N \left(|\mathbf{u}^* \mathbf{D}_j^{-1} \mathbf{r}_j|^2 + \mathbf{u}^* \mathbf{D}_j^{-1} \mathbf{D}_j^{-1*} \mathbf{u} \right) |\hat{\beta}_j|^2 \right)^\ell + K_\ell N^{-\ell+1} v_n^{-6\ell} \\
=\ & K_\ell N^{-\ell} v_n^{-2\ell} \mathrm{E} \left(\sum_{j=1}^N \mathrm{E}_{(j)} |\mathbf{u}^* \mathbf{D}_j^{-1} \mathbf{s}_j|^2 |\hat{\beta}_j|^2 \right)^\ell + K_\ell N^{-\ell+1} v_n^{-6\ell} \\
\le\ & K_\ell N^{-\ell} v^{-2\ell} \mathrm{E} \left(\left(\sum_{j=1}^N |\mathbf{u}^* \mathbf{D}_j^{-1} \mathbf{s}_j|^2 |\beta_j|^2 \right) + N^{1/3} v_n^{-6} \right)^\ell + K_\ell N^{-\ell+1} v_n^{-6\ell} \\
\le\ & K_\ell N^{-\ell} v^{-2\ell} \mathrm{E} \left(\mathbf{u}^* \mathbf{D}^{-1} \mathbf{C}_n \mathbf{D}^{-1*} \mathbf{u} \right)^\ell + K_\ell N^{-2\ell/3} v_n^{-8\ell} \\
\le\ & K_\ell N^{-2\ell/3} v_n^{-8\ell}.
\end{aligned}
\tag{3.24}
$$

Thus (3.7) is proven.

4 Convergence of (3.4) and (3.5)

In this section we will show

$$
|\mathrm{E} m_n - (1/(zn)) \mathrm{tr}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}| \le \begin{cases} K v_n^{-7} n^{-1} & \text{(4.1a)} \\ K n^{-1}, & \text{(4.1b)} \end{cases}
$$

and

$$|\mathrm{E}\mathbf{u}^*\mathbf{D}^{-1}\mathbf{u} - (1/z)\mathbf{u}^*(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{u}| \;\; \leq \;\; Kn^{-1/3}v_n^{-5}. \tag{4.2}$$

The "a" bound and the quadratic form bound holds uniformly for all $x \in [e, f]$. The "b" bound holds uniformly for all $x \in [a_1, a_2]$.

We begin first with some observations on $(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}$. From (3.1), each eigenvalue of $z(\widehat{\mathbf{A}}_n - \mathbf{I})$ equals

$$\frac{t_i}{1 + c_n \mathrm{E} m_n(z)} - (1 + c_n \mathrm{E} m_n(z))z + (1 - c_n),$$

where t_i is an eigenvalue of $(1/N)\mathbf{R}_n\mathbf{R}_n^*$. We see that the imaginary part of this quantity is less than or equal to $-v_n$. Thus we have

$$\|(z\widehat{\mathbf{A}}_n - z\mathbf{I})^{-1}\| \leq v_n^{-1}.$$

Let

$$\widehat{w}_n(z) = (1 + c_n \mathrm{E} m_n(z))^2 z - (1 + c_n \mathrm{E} m_n(z))(1 - c_n),$$

As mentioned in the introduction, the interval $[a_1, a_2]$ corresponds to two eigenvalues $t_1 < t_2$ of $(1/N)R_nR_n^*$, and that for all n large, $w_n^0(x)$ maps $[a_1, a_2]$ into an interval in $[c, d] \subset (t_1, t_2)$. Let t_n^i denote the i-th smallest eigenvalue of $(1/n)R_nR_n^*$. Then we must have for these n the existence of $\hat{\epsilon} > 0$ for which

$$\inf_{x \in [a_1, a_2], i \leq n} |t_n^i - w_n^0(x)| > \hat{\epsilon},$$

and from Lemma 3.1 we get for all n large

$$\inf_{x \in [a_1, a_2], i \leq n} |t_n^i - \widehat{w}_n(x + iv_n)| > \hat{\epsilon}/2. \tag{4.3}$$

We see that the eigenvalues of $(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}$ are

$$\frac{z(1 + c_n \mathrm{E} m_n(z))}{t_n^i - (1 + c_n \mathrm{E} m_n(z))z(\mathrm{E}\mathbf{m}_n(z) + 1)}$$

$$= \frac{z(1 + c_n \mathrm{E} m_n(z))}{t_n^i - (1 + c_n \mathrm{E} m_n(z))(-(1 - c_n) + (1 + c_n \mathrm{E} m_n(z))z)}.$$

Therefore

$$\|(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\| \le \begin{cases} K v_n^{-1} & \text{(4.4a)} \\ K. & \text{(4.4b)} \end{cases}$$

The "a" bound is valid for all $x \in [e.f]$, the "b" bound for all $x \in [a_1, a_2]$.

Let

$$\hat{\beta} = \hat{\beta}(z) = \frac{1}{1 + c_n \mathrm{E} m_n}.$$

We break down the two quantities in a similar fashion. Write

$$\mathbf{D} - (z\widehat{\mathbf{A}}_n - z\mathbf{I}) = \mathbf{C}_n - z\mathbf{I} - (z\widehat{\mathbf{A}}_n - z\mathbf{I}) = \sum_{j=1}^{N} \mathbf{s}_j \mathbf{s}_j^* - z\widehat{\mathbf{A}}_n.$$

Taking inverses and using Lemma 2.2 we get

$$(z\widehat{\mathbf{A}}_n - z\mathbf{I})^{-1} - \mathbf{D}^{-1} = \sum_{j=1}^{N} \beta_j (z\widehat{\mathbf{A}}_n - z\mathbf{I})^{-1} \mathbf{s}_j \mathbf{s}_j^* \mathbf{D}_j^{-1} - (\widehat{\mathbf{A}}_n - \mathbf{I})^{-1} \widehat{\mathbf{A}}_n \mathbf{D}^{-1}. \tag{4.5}$$

Taking traces and expected values, dividing by n, and using the identity

$$\mathbf{m}_n(z) = -\frac{1}{zN} \sum_{j=1}^{N} \beta_j, \tag{4.6}$$

((3.2) of Dozier and Silverstein (2007)) we get

$$(1/n)\mathrm{tr}(z\widehat{\mathbf{A}}_n - z\mathbf{I})^{-1} - \mathrm{E} m_n(z)$$

$$= \sum_{j=1}^{N} \mathrm{E}(z^{-1} \beta_j (1/n) \mathbf{s}_j^* \mathbf{D}_j^{-1} (\widehat{\mathbf{A}}_n - I)^{-1} \mathbf{s}_j) - (1/n)\mathrm{tr}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1} \widehat{\mathbf{A}}_n \mathrm{E}(\mathbf{D}^{-1}).$$

$$:= \frac{1}{zN} \sum_{j=1}^{N} (d_j + e_j + f_j)),$$

where

$$\begin{aligned} d_j &= \frac{1}{n} \mathrm{E}\Big(\beta_j (\mathbf{x}_j^* \mathbf{D}_j^{-1} (\widehat{\mathbf{A}}_n - \mathbf{I})^{-1} \mathbf{x}_j - \mathrm{tr}\mathrm{E}(\mathbf{D}^{-1})(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1})\Big), \\ e_j &= \frac{1}{n} \mathrm{E}\Big(\beta_j (\mathbf{r}_j^* \mathbf{D}_j^{-1} (\widehat{\mathbf{A}}_n - \mathbf{I})^{-1} \mathbf{x}_j + \mathbf{x}_j^* \mathbf{D}_j^{-1} (\widehat{\mathbf{A}}_n - \mathbf{I})^{-1} \mathbf{r}_j)\Big), \\ f_j &= \frac{1}{n} \Big(\mathrm{E}(\beta_j \mathbf{r}_j^* \mathbf{D}_j^{-1} (\widehat{\mathbf{A}}_n - \mathbf{I})^{-1} \mathbf{r}_j) - \hat{\beta} \mathbf{r}_j^* \mathrm{E}(\mathbf{D}^{-1})(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1} \mathbf{r}_j\Big). \end{aligned}$$

In (4.5), we take expected value and pre and post multiply by $\mathbf{u}^*$ and $\mathbf{u}$ to get

$$\begin{aligned}
&\frac{1}{z}\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{u}-\mathrm{E}\mathbf{u}^*\mathbf{D}^{-1}\mathbf{u}\\
=\ &\frac{1}{z}\sum_{j=1}^{N}\mathrm{E}\beta_j\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{u}-\mathrm{E}\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\widehat{\mathbf{A}}\mathbf{D}^{-1}\mathbf{u}\\
:=\ &\frac{1}{z}\sum_{j=1}^{N}(\breve{d}_j+\breve{e}_j+\breve{f}_j),
\end{aligned}$$

where

$$\begin{aligned}
\breve{d}_j &= \frac{1}{N}\mathrm{E}\Big(\beta_j(\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{x}_j\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{u}-\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathrm{E}\mathbf{D}^{-1}\mathbf{u})\Big),\\
\breve{e}_j &= \frac{1}{N}\mathrm{E}\Big(\beta_j(\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{u}+\mathrm{E}\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{x}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u})\Big),\\
\breve{f}_j &= \frac{1}{N}\Big(\mathrm{E}\beta_j\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}-\mathrm{E}\hat{\beta}\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}^{-1}\mathbf{u}\Big).
\end{aligned}$$

We first handle d_j and $\breve{d}_j$. Write $d_j=d_{j1}+d_{j2}+d_{j3}$, where

$$\begin{aligned}
d_{j1} &= \frac{1}{n}\mathrm{E}\beta_j\left(\mathbf{x}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{x}_j-\mathrm{tr}(\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1})\right)\\
d_{j2} &= \frac{1}{n}\mathrm{E}\beta_j\left(\mathrm{tr}(\mathbf{D}_j^{-1}-\mathbf{D}^{-1})(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\right)\\
d_{j3} &= \frac{1}{n}\mathrm{E}\beta_j\left(\mathrm{tr}(\mathbf{D}^{-1}-\mathrm{E}\mathbf{D}^{-1})(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\right).
\end{aligned}$$

We shall show that

$$U_k:=\left|\frac{1}{nz}\sum_{j=1}^{N}d_{jk}\right|\le\begin{cases}Kv_n^{-3} & (4.7\text{a})\\ K. & (4.7\text{b})\end{cases}$$

for $k=1,2,3$.

By Lemma 2.3 and the fact that $\beta_j=\hat{\beta}_j-\beta_j\hat{\beta}_j\gamma_j$, we have

$$U_1\ \le\ \frac{K}{N}\sum_{j=1}^{N}|\mathrm{E}\beta_j\hat{\beta}_j\gamma_j\left(\mathbf{x}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{x}_j-\mathrm{tr}(\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1})\right)|$$

$$
\begin{aligned}
&\le \frac{K}{N}\left(\sum_{j=1}^N \mathrm{E}|\beta_j\gamma_j|^2\right)^{1/2}\left(\sum_{j=1}^N \mathrm{E}|\hat\beta_j^2|\left|\mathbf{x}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{x}_j-\mathrm{tr}(\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1})\right|^2\right)^{1/2}\\
&\le K\left(\sum_{j=1}^N \mathrm{E}|\beta_j\gamma_j|^2\right)^{1/2}\left(\frac{1}{N^2}\sum_{j=1}^N \mathrm{E}|\hat\beta_j^2|\mathrm{tr}((\widehat{\mathbf{A}}_n-\mathbf{I})^{-1*}\mathbf{D}_j^{-1*}\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1})\right)^{1/2}\\
&\le K\left(\sum_{j=1}^N \mathrm{E}|\beta_j\gamma_j|^2\right)^{1/2}\|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\|\left(\frac{1}{N^2}\sum_{j=1}^N \mathrm{E}|\hat\beta_j^2|\mathrm{tr}(\mathbf{D}_j^{-1*}\mathbf{D}_j^{-1})\right)^{1/2}.
\end{aligned}
$$

By Lemma 3.3 we have the last factor

$$
\begin{aligned}
\le\left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}|\hat\beta_j^2|\mathbf{s}_j^*\mathbf{D}_j^{-1*}\mathbf{D}_j^{-1}\mathbf{s}_j\right)^{1/2} &\le \left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}|\beta_j|^2\mathbf{s}_j^*\mathbf{D}_j^{-1*}\mathbf{D}_j^{-1}\mathbf{s}_j\right)^{1/2}+Kn^{-1/3}v_n^{-3}\\
&= \left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}\mathbf{s}_j^*\mathbf{D}^{-1*}\mathbf{D}^{-1}\mathbf{s}_j\right)^{1/2}+Kn^{-1/3}v_n^{-3}
\end{aligned}
$$

Thus, from Corollary 3.1, (3.14) and (4.4) we see that U_1 satisfies (4.7).

We have from Lemma 2.1

$$
\begin{aligned}
U_2 &\le \frac{K}{n}\sum_{j=1}^N |\mathrm{E}\beta_j\left(\mathrm{tr}(\mathbf{D}_j^{-1}-\mathbf{D}^{-1})(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\right)|\\
&= \frac{K}{n}\sum_{j=1}^N |\mathrm{E}\beta_j^2\left(\mathrm{tr}(\mathbf{D}_j^{-1}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}_j^{-1})(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\right)|\\
&\le \frac{K}{N}\|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\sum_{j=1}^N \mathrm{E}|\beta_j^2\mathbf{s}_j^*\mathbf{D}_j^{-2}\mathbf{s}_j| \le \frac{K}{N}\|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\|\sum_{j=1}^N \mathrm{E}|\beta_j^2\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{s}_j|\\
&= \frac{K}{N}\|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\|\sum_{j=1}^N \mathrm{E}\mathbf{s}_j^*\mathbf{D}^{-1*}\mathbf{D}^{-1}\mathbf{s}_j.
\end{aligned}
$$

So we see U_2 satistifes (4.7)

Finally, using (3.14a), (3.1), and (4.6), by martingale decomposition, we have

$$
\begin{aligned}
U_3 &= c_n^{-1}|\mathrm{E}\mathbf{m}_n\mathrm{tr}(\mathbf{D}^{-1}-\mathrm{E}\mathbf{D}^{-1})(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}| = |\mathrm{E}m_n\mathrm{tr}(\mathbf{D}^{-1}-\mathrm{E}\mathbf{D}^{-1})(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}|\\
&= |\mathrm{E}(m_n-\mathrm{E}m_n)\mathrm{tr}(\mathbf{D}^{-1}-\mathrm{E}\mathbf{D}^{-1})(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}|\\
&\le (\mathrm{E}|m_n-\mathrm{E}m_n|^2)^{1/2}(\mathrm{E}|\mathrm{tr}(\mathbf{D}^{-1}-\mathrm{E}\mathbf{D}^{-1})(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}|^2)^{1/2}
\end{aligned}
$$

$$\leq n^{-1}v_n^{-4}\left(\mathrm{E}\Big|\sum_{j=1}^{N}\mathrm{tr}(\mathrm{E}_j\mathbf{D}^{-1}-\mathrm{E}_{j-1}\mathbf{D}^{-1})(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\Big|^2\right)^{1/2}$$

$$= n^{-1}v_n^{-4}\left(\mathrm{E}\Big|\sum_{j=1}^{N}(\mathrm{E}_j-\mathrm{E}_{j-1})\beta_j\mathbf{s}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{D}_j^{-1}\mathbf{s}_j\Big|^2\right)^{1/2}$$

$$\leq n^{-1}v_n^{-4}\left(\sum_{j=1}^{N}\mathrm{E}|\beta_j\mathbf{s}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{D}_j^{-1}\mathbf{s}_j|^2\right)^{1/2} \leq Kn^{-1/2}v_n^{-8}$$

Thus, (4.7) is proven.

Split $\breve{d}_j = \breve{d}_{j1} + \breve{d}_{j2} + \breve{d}_{j3}$, where

$$\begin{aligned}
\breve{d}_{j1} &= \frac{1}{n}\mathrm{E}\beta_j\left(\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{x}_j\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{u} - \mathrm{E}\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{D}_j^{-1}\mathbf{u}\right)\\
\breve{d}_{j2} &= \frac{1}{n}\mathrm{E}\beta_j\left(\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}(\mathbf{D}_j^{-1}-\mathbf{D}^{-1})\mathbf{u}\right)\\
\breve{d}_{j3} &= \frac{1}{n}\mathrm{E}\beta_j\left(\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}(\mathbf{D}^{-1}-\mathrm{E}\mathbf{D}^{-1})\mathbf{u}\right).
\end{aligned}$$

We shall show that

$$\breve{U}_k := \frac{1}{|z|}\left|\sum_{j=1}^{N}\breve{d}_{jk}\right| \leq Kn^{-1/3}v_n^{-5}, \tag{4.8}$$

for $k = 1, 2, 3$.

First, by performing the same steps taken for U_1 and using (3.14a) and (4.4a), we have

$$\begin{aligned}
\breve{U}_1 &\leq \frac{K}{Nv_n}\left(\sum_{j=1}^{N}\mathrm{E}|\hat{\beta}_j^2|\mathbf{u}^*(\widehat{\mathbf{A}}-\mathbf{I})^{-1}(\widehat{\mathbf{A}}-\mathbf{I})^{-1*}\mathbf{D}_j^{-1*}\mathbf{u}\mathbf{u}^*\mathbf{D}_j^{-1*}\mathbf{D}_j^{-1*}\mathbf{u}\right)^{1/2}\\
&\leq \frac{K}{Nv^2}\left(\sum_{j=1}^{N}\mathrm{E}|\hat{\beta}_j^2|\mathbf{u}^*\mathbf{D}_j^{-1*}\mathbf{D}_j^{-1}\mathbf{u}\right)^{1/2} \leq \frac{K}{v_n^2\sqrt{N}}\left(\sum_{j=1}^{N}\mathrm{E}|\hat{\beta}_j^2|\mathbf{u}^*\mathbf{D}_j^{-1*}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{u}\right)^{1/2}\\
&\leq \frac{K}{v_n^2\sqrt{N}}\left(\sum_{j=1}^{N}\mathrm{E}\mathbf{u}^*\mathbf{D}^{-1*}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{u}\right)^{1/2} + Kn^{-1/3}v_n^{-5}\\
&= \frac{K}{v^2\sqrt{N}}\left(\mathrm{E}\mathbf{u}^*\mathbf{D}^{-1}\mathbf{C}_n\mathbf{D}^{-1*}\mathbf{u}\right)^{1/2} + Kn^{-1/3}v_n^{-5} \leq Kn^{-1/3}v_n^{-5}.
\end{aligned}$$

Secondly, we have

$$
\begin{aligned}
\breve{U}_2 &\leq \frac{K}{N}\sum_{j=1}^{N}|\mathrm{E}\beta_j^2\left(\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{D}_j^{-1}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{u}\right)| \\
&\leq \frac{K}{N}\left(\sum_{j=1}^{N}\mathrm{E}|\beta_j\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{D}_j^{-1}\mathbf{s}_j|^2\right)^{1/2}\left(\sum_{j=1}^{N}\mathrm{E}|\beta_j\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{u}|^2\right)^{1/2} \\
&\leq \frac{K}{v_nN}\left(\sum_{j=1}^{N}\mathrm{E}\mathbf{s}_j^*\mathbf{D}^{-1*}\mathbf{D}^{-1}\mathbf{s}_j\right)^{1/2}\left(\sum_{j=1}^{N}\mathrm{E}|\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{u}|^2\right)^{1/2} \\
&= \frac{K}{v_nN}\left(\mathrm{Etr}\mathbf{D}^{-1}\mathbf{C}_n\mathbf{D}^{-1*}\right)^{1/2}\left(\mathrm{E}\mathbf{u}^*\mathbf{D}^{-1*}\mathbf{C}_n\mathbf{D}^{-1}\mathbf{u}\right)^{1/2} \\
&\leq \frac{K}{v_n^3\sqrt{N}}.
\end{aligned}
$$

Finally, similar to the steps taken for U_3 we have

$$
\begin{aligned}
\breve{U}_3 &\leq \left(\mathrm{E}|m_n-\mathrm{E}m_n|^2\mathrm{E}|\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}(\mathbf{D}^{-1}-\mathrm{E}\mathbf{D}^{-1})\mathbf{u}|^2\right)^{1/2} \\
&\leq KN^{-1}v_n^{-6}.
\end{aligned}
$$

Therefore, (4.8) is proven.

We turn now to e_j and $\breve{e}_j$.

Similar to the approach taken for U_1, we have

$$
\begin{aligned}
J &:= \left|\frac{1}{z}\sum_{j=1}^{N}e_j\right| \\
&= \left|\frac{1}{zn}\sum_{j=1}^{N}\mathrm{E}\left(\hat{\beta}_j\beta_j\gamma_j(\mathbf{r}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{x}_j+\mathbf{x}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j)\right)\right| \\
&\leq K\left(\sum_{j=1}^{N}\mathrm{E}|\beta_j\gamma_j|^2\right)^{1/2}\|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\|\left(\frac{1}{N^2}\sum_{j=1}^{N}\mathrm{E}|\hat{\beta}_j|^2\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{r}_j\right)^{1/2},
\end{aligned}
$$

with the last factor

$$
\leq \left(\frac{1}{N}\sum_{j=1}^{N}\mathrm{E}|\hat{\beta}_j^2|\mathbf{s}_j\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{s}_j\right)^{1/2} \leq \left(\frac{1}{N}\sum_{j=1}^{N}\mathrm{E}|\beta_j^2|\mathbf{s}_j\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{s}_j\right)^{1/2} + Kn^{-1/3}v_n^{-3}
$$

$$= \left(\frac{1}{N}\sum_{j=1}^{N}\mathrm{E}\mathbf{s}_j\mathbf{D}^{-1}\mathbf{D}^{-1*}\mathbf{s}_j\right)^{1/2} + Kn^{-1/3}v_n^{-3},$$

so, from Corollary 3.1, (3.14), and (4.4) we have

$$J \leq \begin{cases} Kv_n^{-3} & \text{(4.9a)} \\ K. & \text{(4.9b)} \end{cases}$$

For $\breve{e}_j$ we proceed in a similar fashion. we have by (3.14a) and (4.4a)

$$\begin{aligned}
\breve{J} &:= \left|\frac{1}{z}\sum_{j=1}^{N}\breve{e}_j\right| \\
&\leq K\left(\sum_{j=1}^{N}\mathrm{E}|\hat{\beta}_j^2||\gamma_j^2|\right)^{1/2}\left(\frac{1}{N^2}\sum_{j=1}^{N}\mathrm{E}|\mathbf{u}_j^*\mathbf{D}_j^{-1*}\mathbf{D}_j^{-1}\mathbf{u}|\mathbf{u}^*(\hat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j|^2|\hat{\beta}_j^2|\right. \\
&\qquad \left.+\frac{1}{N^2}\sum_{j=1}^{N}\mathrm{E}|\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}|^2\mathbf{u}^*(\hat{\mathbf{A}}_n-\mathbf{I})^{-1}(\hat{\mathbf{A}}_n-\mathbf{I})^{-1*}\mathbf{u}|\hat{\beta}_j^2|\right)^{1/2} \\
&\leq Kv_n^{-1}\left(\frac{1}{N^2}\sum_{j=1}^{N}v_n^{-4}\mathrm{E}|\mathbf{u}^*(\hat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j|^2 + v_n^{-2}|\beta_j|^2|\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}|^2\right)^{1/2} + Kn^{-1/3}v_n^{-5} \\
&= \frac{K}{\sqrt{N}v_n}\mathrm{E}\left(v_n^{-4}\mathbf{u}^*(\hat{\mathbf{A}}_n-\mathbf{I})^{-1}(1/N)\mathbf{R}_n\mathbf{R}_n^*(\hat{\mathbf{A}}_n-\mathbf{I})^{-1*}\mathbf{u}+v_n^{-2}\mathbf{u}^*\mathbf{D}^{-1}\mathbf{R}_n\mathbf{R}_n^*\mathbf{D}^{-1*}\mathbf{u}\right)^{1/2} \\
&\qquad + Kn^{-1/3}v_n^{-5} \\
&\leq Kn^{-1/3}v_n^{-5}
\end{aligned}$$

Before continuing with f_j and $\breve{f}_j$, we make some observations on $\hat{\beta}$. It is clear that $\hat{\beta} \leq |z|/v_n$. From Lemma 2.1 (c) of Dozier and Silverstein (2007b) we have $\Re(1 + \sigma^2 cm^0(x)) > 0$ for all $x \in \mathbb{R} - \{0\}$. Therefore,

$$K_1 \equiv \inf_{x\in[a_1,a_2]} \Re(1+\sigma^2 cm^0(x)) > 0,$$

and from Lemma 3.1 and the uniform convergence of m_n^0 to m^0 on bounded subsets of $\mathbb{C} \cup [a_1, a_2]$, we have for all n large

$$\inf_{x\in[a_1,a_2]} \Re(1+c_n\mathrm{E}m_n(x+iv_n)) \geq K_1/2.$$

Moreover, since m_n^0 is bounded for $x \in [a_1, a_2]$, we get from Lemma 3.1 that $\hat{\beta} = 1 + c_n \mathrm{E} m_n$ is bounded. Therefore,

$$\sup_{x \in [a_1, a_2]} \max(|\hat{\beta}|, \hat{\beta}^{-1}) \leq \begin{cases} K v_n^{-1} & \text{(4.10a)} \\ K. & \text{(4.10b)} \end{cases}$$

By Lemma 2.2

$$\begin{aligned} f_j &= \frac{\hat{\beta}}{Nn} \mathrm{E}\beta_j \Big(([\mathrm{E}\mathrm{tr}\mathbf{D}^{-1}] - \mathbf{x}_j^* \mathbf{D}_j^{-1}(\mathbf{x}_j + \mathbf{r}_j))\mathbf{r}_j^* \mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j \\ &\quad + (\mathbf{r}_j^* \mathbf{D}_j^{-1}(\mathbf{x}_j + \mathbf{r}_j))\mathbf{x}_j^* \mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j \Big) \end{aligned}$$

Our goal is to show

$$\left| \sum_{j=1}^N f_j \right| \leq \begin{cases} K v_n^{-7} & \text{(4.11a)} \\ K. & \text{(4.11b)} \end{cases}$$

Write

$$\hat{\beta}^{-1} f_j = f_{j1} + f_{j2} + f_{j3} - f_{j4} + f_{j5} + f_{j6} \tag{4.12}$$

where

$$\begin{aligned} f_{j1} &= \frac{1}{N} \mathrm{E}\beta_j (\mathrm{E}m_n(z) - m_n(z))\mathbf{r}_j^* \mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j \\ f_{j2} &= \frac{1}{Nn} \mathrm{E}\beta_j (\mathrm{tr}\mathbf{D}^{-1} - \mathrm{tr}\mathbf{D}_j^{-1})\mathbf{r}_j^* \mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j \\ \hat{f}_{j3} &= \frac{1}{Nn} \mathrm{E}\beta_j (\mathrm{tr}\mathbf{D}_j^{-1} - \mathbf{x}_j^* \mathbf{D}_j^{-1}\mathbf{x}_j)\mathbf{r}_j^* \mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j \\ f_{j4} &= \frac{1}{Nn} \mathrm{E}\beta_j \mathbf{x}_j^* \mathbf{D}_j^{-1}\mathbf{r}_j \mathbf{r}_j^* \mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j \\ f_{j5} &= \frac{1}{Nn} \mathrm{E}\beta_j \mathbf{r}_j^* \mathbf{D}_j^{-1}\mathbf{r}_j \mathbf{x}_j^* \mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j \\ f_{j6} &= \frac{1}{Nn} \mathrm{E}\beta_j \mathbf{r}_j^* \mathbf{D}_j^{-1}\mathbf{x}_j \mathbf{x}_j^* \mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j. \end{aligned}$$

We shall prove

$$\left| \sum_{j=1}^N f_{jk} \right| \leq \begin{cases} K v_n^{-6} & \text{(4.13a)} \\ K, & \text{(4.13b)} \end{cases}$$

for $k = 1, \ldots, 6$.

For $k = 1$

$$
\begin{aligned}
&\frac{1}{N}\sum_{j=1}^{N} |\mathrm{E}\beta_j(\mathrm{E}m_n(z) - m_n(z))\mathbf{r}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j| \\
\leq\ & (\mathrm{E}|m_n - \mathrm{E}m_n|^2)^{1/2}\left(\frac{1}{N}\sum_{j=1}^{N}\left(\mathrm{E}|\beta_j\mathbf{r}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j|^2\right)^{1/2}\right).
\end{aligned}
$$

The last factor is, using Lemma 3.3

$$
\begin{aligned}
\leq\ & \left(\frac{1}{N^2}\sum_{j=1}^{N}\mathrm{E}|\beta_j|^2\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{r}_j\sum_{j=1}^{N}\mathbf{r}_j^*(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1*}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j\right)^{1/2} \\
=\ & \left(\sum_{j=1}^{N}\mathrm{E}|\beta_j|^2\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{r}_j\frac{1}{N}\mathrm{tr}(1/N)\mathbf{R}_n\mathbf{R}_n^*(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1*}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\right)^{1/2} \\
\leq\ & \|(\widehat{A}_n - \mathbf{I})^{-1}\|\left(N\mathrm{E}\sum_{j=1}^{N}|\hat{\beta}_j|^2\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{s}_j + KN^{1/3}v_n^{-6}\right)^{1/2} \\
\leq\ & \|(\widehat{A}_n - \mathbf{I})^{-1}\|\left(N\mathrm{E}\sum_{j=1}^{N}\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{D}^{-1*}\mathbf{s}_j + KN^{4/3}v_n^{-6}\right)^{1/2}.
\end{aligned}
$$

Therefore, from (3.6), Corollary 3.1 and (4.4), we see that (4.13) holds for $k = 1$.

For $k = 2$, using the fact that $\|\mathbf{r}_j\|/\sqrt{N}$ is bounded, we have

$$
\begin{aligned}
&\frac{1}{Nn}\sum_{j=1}^{N}|\mathrm{E}\beta_j(\mathrm{tr}\mathbf{D}^{-1} - \mathrm{tr}\mathbf{D}_j^{-1})\mathbf{r}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j| \\
=\ & \frac{1}{Nn}\sum_{j=1}^{N}|\mathrm{E}\beta_j^2\mathbf{s}_j^*\mathbf{D}_j^{-2}\mathbf{s}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j| \\
=\ & \frac{1}{Nn}\sum_{j=1}^{N}\mathrm{E}|\mathbf{s}_j^*\mathbf{D}^{-2}\mathbf{s}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j| \\
\leq\ & \frac{K}{Nn}\left(\sum_{j=1}^{N}\mathrm{E}|\mathbf{s}_j^*\mathbf{D}^{-2}\mathbf{s}_j|^2\right)^{1/2}\left(\sum_{j=1}^{N}\mathrm{E}|\mathbf{r}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j|^2\right)^{1/2}
\end{aligned}
$$

$$\le \frac{K}{N^{3/2}}\|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\|\left(\sum_{j=1}^N \mathrm{E}|\mathbf{s}_j^*\mathbf{D}^{-2}\mathbf{s}_j|^2\right)^{1/2}\left(\sum_{j=1}^N \mathrm{E}|\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{r}_j|\right)^{1/2}$$

$$\le \frac{K}{N}\|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\|\left(\sum_{j=1}^N \mathrm{E}|\mathbf{s}_j^*\mathbf{D}^{-2}\mathbf{s}_j|^2\right)^{1/2}\left(\sum_{j=1}^N \mathrm{E}|\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{D}^{-1*}\mathbf{s}_j||\beta_j^{-2}|\right)^{1/2}$$

$$\le K\|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\|\left(N^{-1}\sum_{j=1}^N \mathrm{E}|\mathbf{s}_j^*\mathbf{D}^{-2}\mathbf{s}_j|^2\right)^{1/2}$$

$$\times\left(N^{-1}\sum_{j=1}^N \mathrm{E}|\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{D}^{-1*}\mathbf{s}_j|^2\right)^{1/4}\left(N^{-1}\sum_{j=1}^N \mathrm{E}|\beta_j|^{-4}\right)^{1/4}.$$

Therefore, from Corollary 3.1, Lemma 3.2, and (4.4), we see that (4.13) holds for $k=2$

For $k=3$ we have, using Lemma 2.3

$$\frac{1}{Nn}\left|\sum_{j=1}^N \mathrm{E}\beta_j(\mathrm{tr}\mathbf{D}_j^{-1}-\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{x}_j)\mathbf{r}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\right|$$

$$=\frac{1}{Nn}\left|\sum_{j=1}^N \mathrm{E}\beta_j\hat\beta_j\gamma_j(\mathrm{tr}\mathbf{D}_j^{-1}-\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{x}_j)\mathbf{r}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\right|$$

$$\le K\left(\sum_{j=1}^N \mathrm{E}|\beta_j|^2|\gamma_j|^2\right)^{1/2}\left(\frac{1}{N^4}\sum_{j=1}^N \mathrm{E}|\hat\beta_j|^2\mathrm{tr}\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}|\mathbf{r}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j|^2\right)^{1/2}.$$

Again, since $\|\mathbf{r}_j\|/\sqrt{N}$ is bounded, the last factor is

$$\le \|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\|\left(\frac{1}{N^3}\sum_{j=1}^N \mathrm{E}|\hat\beta_j|^2\mathrm{tr}\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{r}_j\right)^{1/2}$$

$$\le \|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\|\left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}|\hat\beta_j|^2(\mathrm{E}_{(j)}\mathbf{s}_j\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{s}_j)^2\right)^{1/2}$$

$$\le \|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\|\left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}|\beta_j|^2(\mathbf{s}_j\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{s}_j)^2+Kn^{-2/3}v_n^{-8}\right)^{1/2}$$

$$\le \|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\|\left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}|\beta_j|^{-4}\right)^{1/4}\left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}(\mathbf{s}_j\mathbf{D}^{-1}\mathbf{D}^{-1*}\mathbf{s}_j)^4\right)^{1/4}+o(1) \quad (4.14)$$

Therefore, from Corollary 3.1, (3.14), (4.4), and Lemma 3.2 we get (4.13) for $k = 3$.

For $k = 4$ we have

$$\begin{aligned}
&\frac{1}{Nn}\left|\sum_{j=1}^{N} \mathrm{E}\beta_j \mathbf{x}_j^* \mathbf{D}_j^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j\right| \\
= \; & \frac{K}{Nn}\left|\sum_{j=1}^{N} |\mathrm{E}\beta_j\hat{\beta}_j\gamma_j \mathbf{x}_j^* \mathbf{D}_j^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j\right| \\
\leq \; & K\left(\sum_{j=1}^{N}\mathrm{E}|\beta_j\gamma_j|^2\right)^{1/2}\left(\frac{1}{N^4}\sum_{j=1}^{N}\mathrm{E}|\hat{\beta}_j^2\mathbf{r}_j^*\mathbf{D}_j^{-1*}\mathbf{D}_j^{-1}\mathbf{r}_j||\mathbf{r}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j|^2\right)^{1/2}.
\end{aligned}$$

We see that the last factor is also bounded by (4.14), so we have (4.13) for $k = 4$.

For $k = 5$ we have

$$\begin{aligned}
&\frac{1}{Nn}\left|\sum_{j=1} \mathrm{E}\beta_j \mathbf{r}_j^* \mathbf{D}_j^{-1}\mathbf{r}_j\mathbf{x}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j\right| \\
= \; & \frac{1}{Nn}\left|\sum_{j=1}^{N} \mathrm{E}\beta_j\hat{\beta}_j\gamma_j \mathbf{r}_j^* \mathbf{D}_j^{-1}\mathbf{r}_j\mathbf{x}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j\right| \\
\leq \; & K\left(\sum_{j=1}^{N}\mathrm{E}|\beta_j\gamma_j|^2\right)^{1/2}\left(\frac{1}{N^4}\sum_{j=1}^{N}\mathrm{E}|\hat{\beta}_j|^2|\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{r}_j|^2\mathbf{r}_j^*(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1*}\mathbf{D}_j^{-1*}\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j\right)^{1/2},
\end{aligned}$$

the last factor (again using the fact that $\|\mathbf{r}_j\|/\sqrt{N}$ is bounded)

$$\begin{aligned}
\leq \; & \left(\frac{1}{N^3}\sum_{j=1}^{N}\mathrm{E}|\hat{\beta}_j|^4(\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{r}_j)^2\right)^{1/4}\left(\frac{1}{N^3}\sum_{j=1}^{N}\mathrm{E}(\mathbf{r}_j^*(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1*}\mathbf{D}_j^{-1*}\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j)^2\right)^{1/4} \\
\leq \; & \left(\frac{1}{N}\sum_{j=1}^{N}\mathrm{E}|\hat{\beta}_j|^4(\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{s}_j)^2\right)^{1/4}\left(\frac{1}{N}\sum_{j=1}^{N}\mathrm{E}(\mathbf{s}_j^*(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1*}\mathbf{D}_j^{-1*}\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{s}_j)^2\right)^{1/4} \\
\leq \; & \left(\frac{1}{N}\sum_{j=1}^{N}\mathrm{E}(\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{D}^{-1*}\mathbf{s}_j)^2 + Kn^{-4/3}v_n^{-12}\right)^{1/4}\left(\frac{1}{N}\sum_{j=1}^{N}\mathrm{E}(\mathbf{s}_j^*(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1*}\mathbf{D}_j^{-1*}\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{s}_j)^2\right)^{1/4}
\end{aligned}$$

For the first factor we immediately use Corollary 3.1 For the second factor we have, using $\mathbf{D}_j^{-1} = \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-1}\beta_j^{-1}$,

$$\begin{aligned}
&\sum_{j=1}^N \mathrm{E}(\mathbf{s}_j^*(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1*}\mathbf{D}_j^{-1*}\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{s}_j)^2 \\
\leq\ & 2\sum_{j=1}^N \mathrm{E}(\mathbf{s}_j^*(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1*}\mathbf{D}^{-1*}\mathbf{D}^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{s}_j)^2 \\
&\quad +2\sum_{j=1}^N \mathrm{E}(\mathbf{s}_j^*(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1*}\mathbf{D}^{-1*}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-1*}\mathbf{D}^{-1}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{s}_j)^2)|\beta_j^{-2}| \\
\leq\ & 2\mathrm{Etr}((\widehat{\mathbf{A}}_n - \mathbf{I})^{-1*}\mathbf{D}^{-1*}\mathbf{D}^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{C}_n)^2 \\
+\ & \sqrt{2}\left(\sum_{j=1}^N \mathrm{E}(\mathbf{s}_j^*\mathbf{D}^{-1*}\mathbf{D}^{-1}\mathbf{s}_j)^8 + \mathrm{Etr}((\widehat{\mathbf{A}}_n - \mathbf{I})^{-1*}\mathbf{D}^{-1*}\mathbf{C}_n\mathbf{D}^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{C}_n)^8\right)^{1/2}\left(\sum_{j=1}^N |\beta_j|^{-4}\right)^{1/2}.
\end{aligned}$$

In the second term, the first term in the first factor is covered by Corollary 3.1. Using Lemma 2.11 we find

$$\begin{aligned}
\mathrm{tr}((\widehat{\mathbf{A}}_n - \mathbf{I})^{-1*}\mathbf{D}^{-1*}\mathbf{D}^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{C}_n)^2 &= \mathrm{tr}((\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{C}_n(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1*}\mathbf{D}^{-1*}\mathbf{D}^{-1})^2 \\
&\leq \|(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\|^4\|\mathbf{C}_n\|^2\mathrm{tr}\mathbf{D}^{-2}\mathbf{D}^{-2*}
\end{aligned}$$

and

$$\begin{aligned}
&((\widehat{\mathbf{A}}_n - \mathbf{I})^{-1*}\mathbf{D}^{-1*}\mathbf{C}_n\mathbf{D}^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{C}_n)^8 = ((\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{C}_n(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1*}\mathbf{D}^{-1*}\mathbf{C}_n\mathbf{D}^{-1})^8 \\
\leq\ & \|(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\|^{16}\|\mathbf{C}_n\|^8\mathrm{tr}(\mathbf{D}^{-1*}\mathbf{C}_n\mathbf{D}^{-1})^8 \leq \|(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\|^{16}\|\mathbf{C}_n\|^{16}\mathrm{tr}(\mathbf{D}^{-8}\mathbf{D}^{-8*}).
\end{aligned}$$

Since $\|\mathbf{C}_n\| \leq KN$ we have using Lemma 2.6 and Corollary 3.1, for B sufficiently large

$$\frac{1}{N}\mathrm{E}\|\mathbf{C}_n\|^2\mathrm{tr}\mathbf{D}^{-2}\mathbf{D}^{-2*} \leq Kv_n^{-4}n\mathrm{P}(\|\mathbf{C}_n\| > B) + \frac{1}{N}B^2\mathrm{Etr}\mathbf{D}^{-2}\mathbf{D}^{-2*} \leq K.$$

Similarly,

$$\frac{1}{N}\mathrm{E}\|\mathbf{C}_n\|^{16}\mathrm{tr}\mathbf{D}^{-8}\mathbf{D}^{-8*} \leq K.$$

Thus, using Lemma 3.2, (3.14), and (4.4) we find (4.13) holds for $k = 5$.

For $k = 6$ we have, by (2.1) and Lemma 3.3

$$\frac{1}{Nn}\left|\sum_{j=1}^N \mathrm{E}\beta_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{x}_j\mathbf{x}_j^*\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j\right|$$

$$
\begin{aligned}
\leq\ & \frac{K}{N^2}\sum_{j=1}^N |\mathrm{E}\hat\beta_j \mathbf{r}_j^*\mathbf{D}_j^{-2}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j| \\
& \quad +\frac{K}{N}\sum_{j=1}^N\left(\frac{1}{N^2}\mathrm{E}|\beta_j-\hat\beta_j|^2\mathrm{E}\mathbf{r}_j^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1*}\mathbf{D}_j^{-1*}\mathbf{D}_j^{-1}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{r}_j\right)^{1/2} \\
\leq\ & K\left(\frac{1}{N^2}\sum_{j=1}^N \mathrm{E}|\hat\beta_j|^2\mathbf{r}_j^*\mathbf{D}_j^{-2}\mathbf{D}_j^{-2*}\mathbf{r}_j\right)^{1/2}\left(\frac{1}{N^2}\sum_{j=1}^N \mathbf{r}_j^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1*}(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\right)^{1/2}+Kn^{-1/3}v_n^{-5} \\
\leq\ & K\|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\|\left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}|\hat\beta_j|^2\mathbf{s}_j^*\mathbf{D}_j^{-2}\mathbf{D}_j^{-2*}\mathbf{s}_j\right)^{1/2}+o(1) \\
\leq\ & K\|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\|\left(\sum_{j=1}^N \mathrm{E}\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{D}^{-1*}\mathbf{s}_j\right)^{1/2}+o(1) \\
\leq\ & K\|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\|\left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}\mathbf{s}_j^*\mathbf{D}^{-2}\mathbf{D}^{-2*}\mathbf{s}_j\right)^{1/2} \\
& +K\|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\|\left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}\mathbf{s}_j^*\mathbf{D}^{-2}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{D}^{-1*}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-2*}\mathbf{s}_j|\beta_j^{-2}|\right)^{1/2}+o(1) \\
\leq\ & K\|(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\|\Big[(\frac{1}{N}\sum_{j=1}^N \mathrm{E}\mathbf{s}_j^*\mathbf{D}^{-2}\mathbf{D}^{-2*}\mathbf{s}_j)^{1/2} \\
+\ & \left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}|\beta_j|^{-4}\right)^{1/4}\left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}(\mathbf{s}_j^*\mathbf{D}^{-2}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{D}^{-1*}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-2*}\mathbf{s}_j)^2\right)^{1/4}\Big]+o(1)
\end{aligned}
$$

Therefore, (4.13) holds for $k=6$, and subsequently (4.11) holds, so we conclude that (4.1) is true.

We proceed now in finding a bound for $\breve{f}_j$. Our goal to to show

$$
\sum_{j=1}^N |\breve{f}_j| \leq \frac{K}{\sqrt{N}v_n^7}. \tag{4.15}
$$

As was done for f_j we find, using Lemma 2.2

$$
\breve{f}_j = \frac{\hat\beta}{N^2}\mathrm{E}\beta_j\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\Big(([\mathrm{E}\mathrm{tr}\mathbf{D}^{-1}]-\mathbf{x}_j^*\mathbf{D}_j^{-1}(\mathbf{x}_j+\mathbf{r}_j)\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{r}_j
$$

$$+\mathbf{r}_j\mathbf{D}_j^{-2}(\mathbf{x}_j+\mathbf{r}_j)\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{u}\Big)$$

Write

$$\hat{\beta}^{-1}\breve{f}_j = \breve{f}_{j1}+\breve{f}_{j2}+\breve{f}_{j3}+\breve{f}_{j4}+\breve{f}_{j5}+\breve{f}_{j6} \tag{4.16}$$

where

$$\begin{aligned}
\breve{f}_{j1} &= \frac{c_n}{N}\mathrm{E}\beta_j\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j(\mathrm{E}m_n(z)-m_n(z))\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}\\
\breve{f}_{j2} &= \frac{1}{N^2}\mathrm{E}\beta_j\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j(\mathrm{tr}\mathbf{D}^{-1}-\mathrm{tr}\mathbf{D}_j^{-1})\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}\\
\breve{f}_{j3} &= \frac{1}{N^2}\mathrm{E}\beta_j\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j(\mathrm{tr}\mathbf{D}_j^{-1}-\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{x}_j)\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}\\
\breve{f}_{j4} &= \frac{1}{N^2}\mathrm{E}\beta_j\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}\\
\breve{f}_{j5} &= \frac{1}{N^2}\mathrm{E}\beta_j\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{r}_j\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{u}\\
\breve{f}_{j6} &= \frac{1}{N^2}\mathrm{E}\beta_j\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{x}_j\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{u}.
\end{aligned}$$

We shall prove

$$\left|\sum_{j=1}^N \breve{f}_{jk}\right| \le \frac{K}{\sqrt{N}v_n^6}. \tag{4.17}$$

for $k=1,\ldots,6$.

For $k=1$ we have by (3.6a)

$$\begin{aligned}
&\frac{c_n}{N}\left|\sum_{j=1}^N \mathrm{E}\beta_j\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j(\mathrm{E}m_n(z)-m_n(z))\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}\right|\\
\le\;& \frac{c_n}{N}\sum_{j=1}^N\Big(\mathrm{E}|m_n(z)-\mathrm{E}m_n(z)|^2\mathrm{E}|\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\beta_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}|^2\Big)^{1/2}\\
\le\;& \frac{K}{N^2v_n^5}\sum_{j=1}^N\Big(\mathrm{E}|\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}|^2\Big)^{1/2}\\
\le\;& \frac{K}{N^{3/2}v_n^6}\left(\sum_{j=1}^N\mathrm{E}|\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}|^2\right)^{1/2} \le \frac{K}{Nv_n^6}\left(\sum_{j=1}^N\mathrm{E}\mathbf{u}^*\mathbf{D}_j^{-1*}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{u}\right)^{1/2}
\end{aligned}$$

$$
\begin{aligned}
&\le \frac{K}{Nv_n^7}\left(\sum_{j=1}^N \mathrm{E}|\beta_j|^2\mathbf{u}^*\mathbf{D}_j^{-1*}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{u}\right)^{1/2} \\
&= \frac{K}{Nv_n^7}\left(\sum_{j=1}^N \mathrm{E}\mathbf{u}^*\mathbf{D}^{-1*}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{u}\right)^{1/2} \\
&= \frac{K}{Nv_n^7}\left(\mathbf{u}^*\mathbf{D}^{-1*}\mathbf{C}_n\mathbf{D}^{-1}\mathbf{u}\right)^{1/2} \le \frac{K}{Nv_n^8}.
\end{aligned}
$$

This proves (4.17) for $k = 1$.

For $k = 2$ we have

$$
\begin{aligned}
&\frac{1}{N^2}\left|\sum_{j=1}^N \mathrm{E}\beta_j\mathbf{u}^*(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j(\mathrm{tr}\mathbf{D}^{-1} - \mathrm{tr}\mathbf{D}_j^{-1})\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}\right| \\
&= \frac{1}{N^2}\left|\sum_{j=1}^N \mathrm{E}\mathbf{u}^*(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j\mathbf{s}_j^*\mathbf{D}^{-2}\mathbf{s}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}\right| \\
&\le \frac{K}{N^{3/2}v_n}\left(\sum_{j=1}^N \mathrm{E}|\mathbf{s}_j^*\mathbf{D}^{-2}\mathbf{s}_j|^2\right)^{1/2}\left(\sum_{j=1}^N \mathrm{E}|\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}|^2\right)^{1/2} \\
&\le \frac{K}{Nv_n}\left(\mathrm{Etr}(\mathbf{D}^{-2}\mathbf{C}_n\mathbf{D}^{-2*}\mathbf{C}_n)\right)^{1/2}\left(\sum_{j=1}^N \mathrm{E}\mathbf{u}^*\mathbf{D}_j^{-1*}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{u}\right)^{1/2} \\
&\le \frac{K}{Nv_n}\left(\mathrm{Etr}(\mathbf{D}^{-2}\mathbf{C}_n\mathbf{D}^{-2*}\mathbf{C}_n)\right)^{1/2}\left(\sum_{j=1}^N \mathrm{E}\mathbf{u}^*\mathbf{D}^{-1*}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{u}||\beta_j^{-2}|\right)^{1/2} \\
&\le \frac{K}{Nv_n^2}\left(\mathrm{Etr}(\mathbf{D}^{-2}\mathbf{C}_n\mathbf{D}^{-2*}\mathbf{C}_n)\right)^{1/2}\left(\mathrm{E}\mathbf{u}^*\mathbf{D}^{-1*}\mathbf{C}_n\mathbf{D}^{-1}\mathbf{u}\right)^{1/2} \\
&\le \frac{K}{\sqrt{N}v_n^5}.
\end{aligned}
$$

Therefore, (4.17) holds for $k = 2$.

For $k = 3$ we have, using Lema 2.3 and (3.9)

$$
\frac{1}{N^2}\left|\sum_{j=1}^N \mathrm{E}\beta_j\mathbf{u}^*(\widehat{\mathbf{A}}_n - \mathbf{I})^{-1}\mathbf{r}_j(\mathrm{tr}\mathbf{D}_j^{-1} - \mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{x}_j)\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}\right|
$$

$$
\begin{aligned}
&= \frac{1}{N^2}\left|\sum_{j=1}^N \mathrm{E}\beta_j\hat\beta_j\gamma_j \mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j(\mathrm{tr}\mathbf{D}_j^{-1}-\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{x}_j)\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}\right| \\
&\le \frac{K}{N^3v_n^3}\left(\sum_{j=1}^N \mathrm{E}\mathrm{tr}(\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*})\right)^{1/2} \\
&\qquad \times\left(\sum_{j=1}^N \mathrm{E}|\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}|^2\right)^{1/2} \\
&\le \frac{K}{Nv_n^5}\left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}|\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}|^2\right)^{1/2} \le \frac{1}{Nv_n^5}\left(\sum_{j=1}^N \mathrm{E}\mathbf{u}^*\mathbf{D}_j^{-1}\mathbf{s}_j\mathbf{s}_j^*\mathbf{D}_j^{-1*}\mathbf{u}\right)^{1/2} \\
&\le \frac{K}{Nv_n^6}\left(\mathrm{E}\mathbf{u}^*\mathbf{D}^{-1*}\mathbf{C}_n\mathbf{D}^{-1}\mathbf{u}\right)^{1/2} \le \frac{K}{Nv_n^7}
\end{aligned}
$$

Thus, we have (4.17) for $k = 3$.

For $k = 4$, again using (3.9), we have

$$
\begin{aligned}
&\frac{1}{N^2}\left|\sum_{j=1}^N \mathrm{E}\beta_j \mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}\right| \\
&= \frac{1}{N^2}\left|\sum_{j=1}^N \mathrm{E}\beta_j\hat\beta_j\gamma_j \mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}\right| \\
&\le \frac{K}{N^3v_n^3}\left(\sum_{j=1}^N \mathrm{E}\mathbf{r}_j^*\mathbf{D}_j^{-1*}\mathbf{D}_j^{-1}\mathbf{r}_j\right)^{1/2}\left(\sum_{j=1}^N \mathrm{E}|\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}|^2\right)^{1/2} \\
&\le \frac{K}{Nv_n^5}\left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}\mathbf{s}_j^*\mathbf{D}^{-1*}\mathbf{D}^{-1}\mathbf{s}_j\right)^{1/2}\left(\frac{1}{N}\sum_{j=1}^N \mathrm{E}|\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{u}|^2\right)^{1/2} \\
&\le \frac{K}{Nv_n^6}\left(N^{-1}\mathrm{tr}\mathbf{D}^{-1*}\mathbf{D}^{-1}\mathbf{C}_n\right)^{1/2}\left(\mathbf{u}^*\mathbf{D}^{-1*}\mathbf{C}_n\mathbf{D}^{-1}\mathbf{u}\right)^{1/2} \le \frac{K}{Nv_n^8},
\end{aligned}
$$

so we see (4.17) is true for $k = 4$.

For $k = 5$ we have

$$
\frac{1}{N^2}\left|\sum_{j=1}^N \mathrm{E}\beta_j \mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{r}_j\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{u}\right|
$$

$$
\begin{aligned}
&= \frac{K}{N^2}\left|\sum_{j=1}^N \mathrm{E}\beta_j\hat\beta_j\gamma_j \mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{r}_j\mathbf{x}^*\mathbf{D}_j^{-1}\mathbf{u}\right| \\
&\le \frac{K}{N^3v_n^3}\left(\sum_{j=1}^N \mathrm{E}\mathbf{u}^*\mathbf{D}_j^{-1*}\mathbf{D}_j^{-1}\mathbf{u}\right)^{1/2}\left(\sum_{j=1}^N \mathrm{E}|\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{r}_j|^2\right)^{1/2} \\
&\le \frac{K}{N^2v_n^5}\left(\sum_{j=1}^N \mathrm{E}|\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{r}_j^*|^2\right)^{1/2} \le \frac{K}{\sqrt{N}v_n^6}.
\end{aligned}
$$

Therefore, we have (4.17) for $k=5$.

Finally, for $k=6$ we have

$$
\begin{aligned}
&\frac{1}{N^2}\left|\sum_{j=1}^N \mathrm{E}\beta_j \mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{x}_j\mathbf{x}_j^*\mathbf{D}_j^{-1}\mathbf{u}\right| \\
&\le \frac{K}{N^2v_n}\left(\sum_{j=1}^N \mathrm{E}\mathbf{r}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{r}_j\right)^{1/2}\left(\sum_{j=1}^N \mathrm{E}\mathbf{u}^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{u}|\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}\mathbf{r}_j|^2\right)^{1/2} \\
&\le \frac{K}{Nv_n^2}\left(\sum_{j=1}^N \mathrm{E}\mathbf{s}_j^*\mathbf{D}_j^{-1}\mathbf{D}_j^{-1*}\mathbf{s}_j\right)^{1/2}\left(\mathbf{u}^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1}(1/N)\mathbf{R}_n\mathbf{R}_n^*(\widehat{\mathbf{A}}_n-\mathbf{I})^{-1*}\mathbf{u}\right)^{1/2} \\
&\le \frac{K}{Nv_n^4}\left(\sum_{j=1}^N \mathrm{E}\mathbf{s}_j^*\mathbf{D}^{-1}\mathbf{D}^{-1*}\mathbf{s}_j\right)^{1/2} = \frac{K}{Nv_n^4}\left(\mathrm{E}\mathrm{tr}\mathbf{D}^{-1}\mathbf{D}^{-1*}\mathbf{C}_n\right)^{1/2} \\
&\le \frac{K}{\sqrt{N}v_n^5}
\end{aligned}
$$

Thus we have (4.17) for $k=6$, so that (4.15) is true, and consequently we have (4.2).

5 Proof of Theorem 1.2

We begin by proving

$$
\sup_{x\in[a_1,a_2]} |\mathrm{E}m_n - m_n^0| \le Kn^{-1}. \tag{5.1}
$$

The following is valid for any $z=x+iv$, with $v\in(0,1]$, which will be useful in the proof of Lemma 3.1. Let $b_n^0 = 1+c_nm_n^0$. Then, from (1.1), $b_n^0=b_n^0(z)$ satisfies, for any

$z \in \mathbb{C}^+$

$$b_n^0 = 1 + c_n \int \frac{1}{\frac{t}{b_n^0} - b_n^0 z + (1-c_n)} dH_n(t)$$

$$= 1 + c_n \int \frac{\frac{t b_n^0}{|b_n^0|^2} - \overline{b_n^0 z} + (1-c_n)}{\left|\frac{t}{b_n^0} - b_n^0 z + (1-c_n)\right|^2} dH_n(t)$$

$$= 1 + g_n^0 b_n^0 + ((1-c_n) - \overline{b_n^0 z}) G_n^0,$$

where

$$g_n^0 = c_n \int \frac{\frac{t}{|b_n^0|^2}}{\left|\frac{t}{b_n^0} - b_n^0 z + (1-c_n)\right|^2} dH_n(t)$$

and

$$G_n^0 = c_n \int \frac{1}{\left|\frac{t}{b_n^0} - b_n^0 z + (1-c_n)\right|^2} dH_n(t).$$

Write

$$b_n^0(1-g_n^0) = 1 + (1-c_n)G_n^0 - \overline{b_n^0 z} G_n^0. \tag{5.2}$$

For the following argument a subscript of "2" will denote the imaginary part of a complex number. We have $0 < b_{n2}^0 \leq c_n/v$ and

$$v \leq (b_n^0 z)_2 = v + c_n v \int \frac{\lambda}{|\lambda - z|^2} dF^{c_n, H_n}(\lambda) \leq \frac{K}{v}. \tag{5.3}$$

Then, since $(b_n^0)_2(1-g_n^0) = (b_n^0 z)_2 G_n^0$, we must have $1 - g_n^0 > 0$.

From (5.2) we have

$$b_n^0 z(1-g_n^0) = z(1 + (1-c_n))G_n^0) - \overline{b_n^0}|z|^2 G_n^0.$$

Plugging this into (5.2) gives us

$$b_n^0(1-g_n^0)^2 = (1+(1-c_n)G_n^0)(1-g_n^0) - G_n^0(\overline{z}(1+(1-c_n)G_n^0) - b_n^0|z|^2 G_n^0).$$

Therefore,

$$b_n^0((1-g_n^0)^2 - (G_n^0|z|)^2) = (1+(1-c_n)G_n^0)(1-g_n^0 - G_n^0\overline{z}). \tag{5.4}$$

It follows that

$$b_n^0 z(1 - g_n^0 - G_n^0|z|) = (1 + (1 - c_n)G_n^0)\frac{(1 - g_n^0)z - G|z|^2}{1 - g_n^0 + G_n^0|z|},$$

which implies

$$(b_n^0 z)_2(1 - g_n^0 - G_n^0|z|) = (1 + (1 - c_n)G_n^0)\frac{(1 - g_n^0)v}{1 - g_n^0 + G_n^0|z|}.$$

Therefore, $1 - g_n^0 > G_n^0|z|$. Moreover,

$$\frac{K}{v}(1 - g_n^0 - G_n^0|z|) \geq (b_n^0 z)_2(1 - g_n^0 - G_n^0|z|) \geq \frac{(1 - g_n^0)v}{1 - g_n^0 + G_n^0|z|} \geq \frac{(1 - g_n^0)v}{2(1 - g_n^0)} = \frac{v}{2}. \quad (5.5)$$

Therefore, for all positive $v \leq 1$, there exists a $\underline{K}$ such that

$$1 - g_n^0 \geq G_n^0|z| + \underline{K}v^2.$$

Therefore,

$$g_n^0 + G_n^0|z| \leq 1 - \underline{K}v^2. \quad (5.6)$$

Let $b_n = 1 + c_n \mathrm{E} m_n$ and $w_n = \mathrm{E} m_n - (1/zn)\mathrm{tr}(\widehat{\mathbf{A}}_n - I)^{-1}$. Then

$$\mathrm{E} m_n = \int \frac{1}{\frac{t}{b_n} - b_n z + 1 - c_n} dH_n(t) + w_n \quad (5.7)$$

Define g_n, G_n to be the analogs of g_n^0 and G_n^0 when b_n^0 is replaced by $b_n \equiv 1 + c_n \mathrm{E} m_n$ in the definitions of g_n^0 and G_n^0. We have $\mathrm{E} m_n - m_n^0 = (\mathrm{E} m - m_n^0)\alpha_n + w_n$ where

$$\alpha_n = c_n \int \frac{\frac{t}{b_n b_n^0} + z}{(\frac{t}{b_n} - b_n z + (1 - c_n))(\frac{t}{b_n^0} - b_n^0 z + (1 - c_n))} dH_n(t). \quad (5.8)$$

Using (5.6) we have for n large

$$|\alpha_n| \leq c_n \int \frac{\frac{t}{|b_n||b_n^0|}}{|\frac{t}{b_n} - b_n z + (1 - c_n)|\,|\frac{t}{b_n^0} - b_n^0 z + (1 - c_n)|} dH_n(t)$$
$$+ c_n|z| \int \frac{1}{|\frac{t}{b_n} - b_n z + (1 - c_n)|\,|\frac{t}{b_n^0} - b_n^0 z + (1 - c_n)|} dH_n(t)$$

$$\leq (g_n)^{1/2}(g_n^0)^{1/2} + |z|(G_n)^{1/2}(G_n^0)^{1/2} \leq (g_n + |z|G_n)^{1/2}(g_n^0 + |z|G_n^0)^{1/2}$$
$$\leq (g_n + |z|G_n)^{1/2}(1 - (\underline{K}/2)v_n^2). \tag{5.9}$$

At this point we assume $x \in [a_1, a_2]$ and $v = v_n$. Because of assumption (d) in Theorem 1.1 and Lemma 2.6, the limiting p.d.f. F has bounded support, which, consequently, is true for all F^{c_n,H_n}, uniformly in n. Therefore we see that the integral in (5.3) is uniformly bounded above for all $x \in [a_1, a_2]$. Thus

$$\sup_{x\in[a_1,a_2]} (b_n^0 z)_2 \leq K v_n,$$

which together with (5.5) implies

$$\sup_{x\in[a_1,a_2]} g_n^0 + G_n^0|z| \leq 1 - \underline{K}, \tag{5.10}$$

for some positive $\underline{K}$. Therefore, from (5.9) we have

$$|\alpha_n| \leq (g_n + |z|G_n)^{1/2}(1 - \underline{K})^{1/2}.$$

From (4.3), (4.10b), and the expression for the eigenvalues of $(\widehat{\mathbf{A}}_n - I)^{-1}$ (below (4.3)), it is clear that both g_n and G_n are continuous functions of $\mathrm{E}m_n$. Therefore, from Lemma 3.1 and (5.10), we have for all n large

$$\sup_{x\in[a_1,a_2]} g_n^0 + G_n^0|z| \leq 1 - \underline{K}/2.$$

Therefore, from (4.1b) we conclude

$$\sup_{x\in[a_1,a_2]} |\mathrm{E}m_n - m_n^0| \leq K|w_n| \leq Kn^{-1}.$$

Combining this with (3.6b) we have for all $\ell \geq 1$

$$n\mathrm{E}|m_n - m_n^0|^\ell \leq K.$$

This bound is uniform for all $x \in [a_1, a_2]$ Let S_n be a set of n numbers, equally spaced in $[a_1, a_2]$. Let ℓ be large enough so that nv_n^ℓ is summable. Then, since $|m_n(x_1 + iv_n) - m_n(x_2 + iv_n)| \leq |x_1 - x_2|v_n^{-2}$, for any $\varepsilon > 0$ we have for all n large

$$\begin{aligned}\mathrm{P}(&\sup_{x\in[a_1,a_2]} |nv_n|m_n(x + iv_n) - m_n^0(x + iv_n)| > \varepsilon) \\ &\leq \mathrm{P}(\max_{x\in S_n} |nv_n|m_n(x + iv_n) - m_n^0(x + iv_n)| > \varepsilon/2) \\ &\quad \leq (2/\varepsilon)^\ell n v_n^\ell,\end{aligned}$$

which is summable. Therefore we have Theorem 1.2.

6 Proof of Lemma 3.1 and Corollary 3.1

We begin with the proof of Lemma 3.1. Let $\lambda_{\max}$ denote the largest eigenvalue of C_n. Let B be a bound on the largest eigenvalue of $(1/N)R_nR_n^*$. Let $K_{\max} > (B^{1/2} + (1+\sqrt{c_n}))^2$ for all n. From Lemma 2.6 we have for any $t > 0$

$$|b_n| \geq b_{n2} \geq c_n \mathrm{E} m_{n2} I(\lambda_{\max} \leq K_{\max}) \geq \frac{c_n v_n}{(x+K_{\max})^2 + v_n^2}(1 - o(n^{-t})) \geq K v_n$$

for all n large. Using this bound we find

$$\left|\frac{t}{b_n} - b_n z + (1-c_n))\right|^2 \leq (Kv_n^{-1} + (1-c_n))^2 \leq K v_n^{-2}.$$

Therefore we have $G_n \geq K v_n^2$.

From (5.7) we have

$$b_n(1-g_n) = 1 + (1-c_n)G_n - \overline{b_n z} G_n + w_n. \tag{6.1}$$

Similar to the way (5.4) is obtained we find

$$b_n((1-g_n)^2 - (G_n|z|)^2) = (1+(1-c_n)G_n)(1-g_n-G_n\overline{z}) - G_n\overline{z w_n} + w_n(1-g_n). \tag{6.2}$$

Since $(b_n z)_2 \geq v_n$ and $v_n^3 - |w_n| \geq v_n^3 - Kn^{-1}v_n^{-7}$, we see from (6.1) that $1 - g_n > 0$ for all n large. For these n we see from (6.2) that

$$b_n(1-g_n-G_n|z|) = (1+(1-c_n)G_n)\frac{(1-g_n)-G\overline{z}}{1-g_n+G_n|z|} + A_n w_n + B_n\overline{w}_n$$

where A_n and $|B_n|$ are both contained in (0,1). We see that

$$\frac{(1+(1-c_n)G_n)G_n v_n}{1-g_n+|z|G_n} + (A_n w_n + B_n \overline{w}_n)_2 \geq |z|^{-1}v_n - 2|w_n| \geq Kv_n - 2n^{-1}v_n^{-7} > 0,$$

from which we can conclude that $1 - g_n > G_n|z|$. Therefore, from (5.9) we get

$$|\mathrm{E}m_n - m_n^0| \leq (w/\underline{K})v_n^{-2}|w_n| \leq Kn^{-1}v_n^9 \tag{6.3}$$

This bound is uniform for all $x \in [e, f]$.

As before we let S_n be a set of n numbers equally spaced in $[e,f]$. From (3.6a) and (6.3) we get for any $\ell > 1$ and $\varepsilon > 0$

$$\mathrm{P}(\sup_{x\in[e.f]} v_n^{-1}|m_n - m_n^0| > \varepsilon) \le \mathrm{P}(\max_{x\in S_n} v_n^{-1}|m_n - m_n^0| > \varepsilon/2)$$
$$\le \mathrm{P}(\max_{x\in S_n} v_n^{-1}|m_n - \mathrm{E}m_n^0| > \varepsilon/4) \le \varepsilon^{-\ell} K_\ell n n^{-\ell} v_n^{-5\ell}.$$

Fix $r > 0$. Then for any $\ell > \max(r,2)$ we get by Lemma 2.8

$$\mathrm{E}(v_n^{-r}|\sup_{x\in[e,f]} |m_n - m_n^0|^r) \le (K_\ell n n^{-\ell} v_n^{-5\ell})^{r/\ell} \frac{\ell}{\ell - r} \to 0. \tag{6.4}$$

as $n \to \infty$. Therefore, we have (b) of Lemma 3.1.

For (a) we notice that all the "a" arguments leading to (6.4) apply if we consider v_n fixed in $(0,1)$. Moreover, it is clear that we can find r and $\ell > r$ for which (6.4) is summable. Therefore we get Lemma 3.1(a).

We proceed with the proof of Corollary 3.1. Let $\underline{\epsilon} \in (0,1)$ be such that $[a_1', a_2'] \subset (b_1, b_2)$ with $a_1' = a - \underline{\epsilon}$, $a_2' = a_2 + \underline{\epsilon}$. Assume the interval $[e,f]$ covers $[a_1', a_2']$.

Write $m_{n2} = \Im m_n = m_{n2}^{\text{out}} + m_{n2}^{\text{in}}$ where

$$m_{n2}^{\text{out}}(x + iv_n) = \frac{1}{n} \sum_{\lambda_j \in [a_1', a_2']} \frac{v_n}{(x-\lambda_j)^2 + v_n^2}.$$

Similarly write $m_{n2}^0 = \Im m_n^0 = m_{02}^{\text{out}} + m_{02}^{\text{in}}$ where

$$m_{02}^{\text{out}}(x + iv_n) = \int_{x\in[a_1', a_2']} \frac{v_n}{(x-t)^2 + v_n^2} dF^{c_n, H_n}(t) = 0.$$

From Lemma 3.1

$$\lim_{n\to\infty} \mathrm{E}(v_n^{-r} \sup_{x\in\mathbb{R}} |m_{n2}(x+iv_n) - m_{n2}^0(x+iv_n)|^r) = 0. \tag{6.5}$$

Notice the family of functions

$$f_x(t) = \frac{1}{(x-t)^2}$$

defined on $[a_1', a_2']^c$ is bounded and equicontinuous for $x \in [a_1, a_2]$. We apply Lemma 2.7 with $S = [a_1', a_2']^c$, $\Theta = [a_1, a_2]$, and identifying P_n with either the measures on $[a_1', a_2']^c$

induced by $G_n \equiv F^{C_n}/F^{C_n}\{[a_1', a_2']^c\}$ or F^{c_n,H_n}, and P the measure on $[a_1', a_2']^c$ induced by $F^{c,H}$. Since, as $n \to \infty$, $F^{C_n} \xrightarrow{\mathcal{D}} F^{c,H}$, a.s., $F^{c_n,H_n} \xrightarrow{\mathcal{D}} F^{c,H}$, and $F^{C_n}\{[a_1', a_2']^c\} \to 1$, a.s. we have

$$\sup_{x\in[a_1,a_2]} v_n^{-1}|m_{n2}^{\text{in}} - m_{02}^{\text{in}}| = \sup_{x\in[a_1,a_2]} \left|\int_{t\in[a'b']^c} \frac{1}{(t-x)^2+v_n^2} d(F^{C_n}(t) - F^{c_n,H_n}(t))\right|$$

$$\leq 2\frac{v_n^2}{\underline{\epsilon}^4} + \frac{1}{\underline{\epsilon}^2}\left(1 - \frac{1}{F^{C_n}([a'b']^c)}\right) + \sup_{x\in[a_1,a_2]} \left|\int_{t\in[a'b']^c} \frac{1}{(t-x)^2} d(G_n(t) - F^{c,H}(t))\right|$$

$$+ \sup_{x\in[a_1,a_2]} \left|\int_{t\in[a'b']^c} \frac{1}{(t-x)^2} d(F^{c_n,H_n}(t) - F^{c,H}(t))\right| \to 0, \quad \text{a.s.}$$

as $n \to \infty$. Therefore, since $\sup_{x\in[a_1,a_2]} v_n^{-1}|m_{n2}^{\text{in}} - m_{02}^{\text{in}}|$ is bounded, we get from dominated convergence for $r > 0$

$$\lim_{n\to\infty} \mathrm{E} \sup_{x\in[a_1,a_2]} v_n^{-r}|m_{n2}^{\text{in}} - m_{02}^{\text{in}}|^r = 0,$$

and with (6.5) we conclude

$$\lim_{n\to\infty} \mathrm{E} \sup_{x\in[a_1,a_2]} v_n^{-r}(m_{n2}^{\text{out}})^r = 0, \tag{6.6}$$

For any $x \in [a_1, a_2]$,

$$v_n^{-1} m_{n2}^{\text{out}}(x + iv_n) \geq \int_{[a_1,a_2]} \frac{1}{(t-x)^2+v_n^2} dF^{C_n}(t)$$

$$\geq \int_{[a_1,a_2]\cap[x-v_n,x+v_n]} \frac{1}{(t-x)^2+v_n^2} dF^{C_n}(t) \geq \frac{1}{2v_n^2} F^{C_n}([a_1,a_2]\cap[x-v_n, x+v_n]).$$

Select $x_j \in [a_1, a_2]$ such that $v_n < x_j - x_{j-1}$ and $\cup_j[x_j - v_n, x_j + v_n] \supset [a_1, a_2]$. Then, for $r > 0$, we have from (6.6)

$$v_n^{-r}\mathrm{E}(F^{C_n}([a_1, a_2]))^r \leq v_n^{-r}\mathrm{E}\Big(\sum_j F^{C_n}([a_1, a_2] \cap [x_j - v_n, x_j + v_n])\Big)^r$$

$$\leq v_n^{-r}\mathrm{E}\Big(2\sum_j (x_j - x_{j-1}) \sup_{x\in[a_1,a_2]} (m_{n2}^{\text{out}})\Big)^r$$

$$\leq 2^r(b-a)^r \sup_{x\in[a_1,a_2]} v_n^{-r}\mathrm{E}(m_{n2}^{\text{out}})^r \to 0,$$

as $n \to \infty$. Thus far $v_n = \kappa n^{-1/(2p)}$, with $p > 24$. We will also need p large enough so that

$$N^{d-1} v_n^{-4} = O(1), \tag{6.7}$$

where d is given in condition (1.10) in order for Lemma 3.2 to hold. So we have for any $r > 0$

$$\mathrm{E}(F^{C_n}([a_1, a_2]))^r = o(v_n^r) = o(n^{-r/p})$$

The same result applies as well to $[a'b']$ and so we have

$$\mathrm{E}(F^{C_n}([a_1', a_2']))^r = o(n^{-r/p}).$$

Select now $v_n = \kappa n^{-1/(2mp)}$, for which integer $m \geq 8$. We get then

$$\mathrm{E}(F^{C_n}([a_1', a_2']))^r = o(v_n^{16r}). \tag{6.8}$$

Write $\mathbf{C} = \mathbf{U}\Lambda\mathbf{U}^*$, $\Lambda = \mathrm{diag}[\lambda_1, \cdots, \lambda_n]$ in its spectral decomposition, and for any real $a > 0$,

$$\mathbf{D}_a = \mathbf{U}\mathrm{diag}\left(\frac{1}{|\lambda_1 - z|^a}, \cdots, \frac{1}{|\lambda_p - z|^a}\right)\mathbf{U}^*,$$

Notice for $a \leq 16$, $x \in [a_1, a_2]$, and $r \geq 1$, from (6.8)

$$\begin{aligned} \mathrm{E}\left(\mathrm{tr}\mathbf{D}_a\right)^r &= \mathrm{E}\left(\frac{1}{N}\sum_{\lambda_k \in [a_1', a_2']^c} \frac{1}{|\lambda_k - z|^a} + \frac{1}{N}\sum_{\lambda_k \in [a_1', a_2']} \frac{1}{|\lambda_k - z|^a}\right)^{rl} \\ &\leq K\mathrm{E}(\underline{\epsilon}^{-a} + v_n^{-a} F^{\mathbf{C}_n}[a_1', a_2'])^r \leq K. \end{aligned} \tag{6.9}$$

We get immediately Corollary 3.1(b).

By Jensen,'s inequality we have for any real $m \geq 1$ and $a > 0$

$$\begin{aligned} (\mathbf{s}_j^* \mathbf{D}_a \mathbf{s}_j)^m &= \left(\sum_{k=1}^n \frac{|\mathbf{u}_k^* \mathbf{s}_j|^2}{|\lambda_k - z|^a}\right)^m \leq \left(\sum_{k=1}^n \frac{|\mathbf{u}_k^* \mathbf{s}_j|^2}{|\lambda_k - z|^{am}}\right)\left(\sum_{k=1}^n |\mathbf{u}_k^* \mathbf{s}_j|^2\right)^{m-1} \\ &= \|\mathbf{s}_j\|^{2(m-1)} \mathbf{s}_j^* \mathbf{D}_{am} \mathbf{s}_j \leq K \mathbf{s}_j^* \mathbf{D}_{am} \mathbf{s}_j. \end{aligned}$$

(considering $|\mathbf{u}_k^* \mathbf{s}_j|^2 / \sum_k |\mathbf{u}_k^* \mathbf{s}_j|^2$ as a probability).

Therefore, by Hölder's inequality, (with $\nu = \sum_i(\nu_i + \mu_i)$ and $\iota_k = (\nu_k + \mu_k)/\nu$)

$$
\begin{aligned}
& \mathrm{E}\left|\frac{1}{N}\sum_{j=1}^{N}\prod_{k=1}^{m}\mathbf{s}_j^*\mathbf{D}^{-\nu_k}\mathbf{D}^{-\mu_k *}\mathbf{s}_j\right|^{\ell} \leq \mathrm{E}\left(\frac{1}{N}\sum_{j=1}^{N}\prod_{k=1}^{m}\mathbf{s}_j^*\mathbf{D}_{\nu_k+\mu_k}\mathbf{s}_j\right)^{\ell} \\
\leq \ & \mathrm{E}\left(\prod_{k=1}^{m}\left(\frac{1}{N}\sum_{j=1}^{N}(\mathbf{s}_j^*\mathbf{D}_{\nu_k+\mu_k}\mathbf{s}_j)^{1/\iota_k}\right)^{\iota_k}\right)^{\ell} \leq K\mathrm{E}\left(\frac{1}{N}\sum_{j=1}^{N}\mathbf{s}_j^*\mathbf{D}_{\nu}\mathbf{s}_j\right)^{\ell} \\
= \ & K\mathrm{E}\left(\frac{1}{N}\mathrm{tr}\mathbf{C}_n\mathbf{D}_{\nu}\right)^{\ell}.
\end{aligned}
$$

We have by (6.9)

$$
\mathrm{E}\left(\frac{1}{N}\mathrm{tr}\mathbf{C}_n\mathbf{D}_{\nu}\right)^{\ell} \leq K\mathrm{E}\left(\frac{1}{N}\mathrm{tr}\mathbf{D}_{\nu-1}\right)^{\ell} + K\mathrm{E}\left(\frac{1}{N}\mathrm{tr}\mathbf{D}_{\nu}\right)^{\ell} \leq K.
$$

This gives us Corollary 3.1(a).

7 Proof of Lemma 3.2

We use condition (1.10). Consider a j which satisfies the property in (1.10), that is, no eigenvalues of $(1/(N-1))\mathbf{R}_{nj}\mathbf{R}_{nj}^*$ appear in $(a_1 - \hat{\epsilon}, a_2 + \hat{\epsilon})$. We apply (3.7) and (4.2) to

$$
\mathbf{C}_n^j \equiv \frac{1}{N-1}\sum_{k\neq j}(\mathbf{x}_k + \mathbf{r}_k)(\mathbf{x}_k + \mathbf{r}_k)^*,
$$

where $\mathbf{D}$ is replaced by $\mathbf{D}^j(z) = \mathbf{C}_n^j - zI$ and $\widehat{\mathbf{A}}_n$ is replaced by

$$
\widehat{\mathbf{A}}_{nj} = \frac{1}{z(1 + c_n'\mathrm{E}m_j(z))}(1/(N-1))\mathbf{R}_{nj}\mathbf{R}_{nj}^* - \mathrm{E}\mathbf{m}_{\mathrm{j}}(\mathrm{z})\mathrm{I},
$$

where $\mathbf{R}_{nj}$ is $\mathbf{R}_n$ with the j-th column removed, $m_j(z) = (1/n)\mathrm{tr}{\mathbf{D}^j}^{-1}(z)$, $c_n' = n/(N-1)$, and $\mathbf{m}_{\mathrm{j}}(\mathrm{z})$ is given by (3.1) (c_n replaced by c_n'). We have

$$
\|\mathbf{D}_j^{-1} - {\mathbf{D}^j}^{-1}\| = \|\mathbf{D}_j^{-1}(\mathbf{C}_j - \mathbf{C}_n^j){\mathbf{D}^j}^{-1}\| \leq Kv_n^{-2}N^{-1}. \tag{7.1}
$$

Therefore, using Lemma 2.1 we have

$$
|m_n - m_j| \leq (nv)^{-1} + Kv_n^{-2}N^{-1} \leq Kv_n^{-2}n^{-1}. \tag{7.2}
$$

Recall the definition of $\widehat{w}_n$ and the role of the interval $[c,d]$ in Section 4. Let

$$\widehat{w}_j(z) = (1 + c_n' \mathrm{E} m_j(z))^2 z - (1 + c_n' \mathrm{E} m_j(z))(1 - c_n),$$

The eigenvalues of $(\widehat{\mathbf{A}}_{nj} - \mathbf{I})^{-1}$ are

$$\frac{z(1 + c_n' \mathrm{E} m_j(z))}{t_{nj}^i - (1 + c_n \mathrm{E} m_j(z)) z (\mathrm{E}\mathbf{m}_j(z) + 1)} \\ = \frac{z(1 + c_n \mathrm{E} m_n(z))}{t_{nj}^i - \widehat{w}_j(z)},$$

where t_{nj}^i is the i-th smallest eigenvalue of $1/(N-1)\mathbf{R}_{nj}\mathbf{R}_{nj}^*$. We have all t_{nj}^i lying uniformly away from $[c,d]$. From (7.2) we have $|\widehat{w}_n - \widehat{w}_j| \leq K n^{-1} v_n^{-2}$. Therefore, using (4.3), we find

$$\sup_{x \in [a_1, a_2]} \| |(\widehat{\mathbf{A}}_{nj} - I)^{-1} \| \leq K, \tag{7.3}$$

for all n large. With E^j denoting conditional expectation with respect to $\mathbf{x}_j$, we have using (3.7), (4.2), (7.3), and (7.1) we have for any $\ell \geq 1$ and $x \in [a_1, a_2]$

$$\begin{aligned} \mathrm{E}|\beta_j^{-1}|^\ell &\leq K_\ell (1 + \mathrm{E}_{(j)} (\mathrm{E}^j |\mathbf{s}_j^* \mathbf{D}_j^{-1} \mathbf{s}_j|^\ell)) \\ &\leq K_\ell (1 + \mathrm{E}_{(j)} (\mathrm{E}^j |\mathbf{s}_j^* (\mathbf{D}_j^{-1} - \mathbf{D}^{j^{-1}}) \mathbf{s}_j|^\ell + \mathrm{E}^j |\mathbf{s}_j^* (\mathbf{D}^{j^{-1}} - (1/z)(\widehat{\mathbf{A}}_{nj} - \mathbf{I})^{-1}) \mathbf{s}_j|^\ell) \\ &\leq K_\ell (1 + v_n^{-2\ell} N^{-\ell} + v_n^{5\ell} n^{-\ell/6} + v_n^{-5\ell} n^{-\ell/3}) \leq K_\ell. \end{aligned}$$

Finally, with I_n denoting the index set of j's which satisfy the property in (1.10), we get, using (6.7), for any $\ell \geq 1$

$$\mathrm{E} \left(\frac{1}{N} \sum_{j=1}^N |\beta_j|^{-4} \right)^\ell \leq K_\ell \left((N^{d-1} v_n^{-4})^\ell + \frac{1}{N} \sum_{j \in I_n} \mathrm{E}|\beta_j|^{-4\ell} \right) \leq K_\ell.$$

Therefore, we have Lemma 3.2

References

1. Bai, Z. D. and Silverstein, J. W. (1998) . No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices, *Annals of Probability* **26**, 316-345.

2. Bai, Z.D. and Silverstein, J.W. (1999). Exact separation of eigenvalues of large dimensional sample covariance matrices, *Annals of Probability* 27(3) (1999), pp. 1536-1555.

3. Bai, Z.D. and Silverstein, J.W. (2009). *Spectral Analysis of Large Dimensional Random Matrices.* Springer, New York.

4. Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

5. Capitaine, M., Donati-Martin, C., and Féral, D. (2009). The largest eigenvalue of finite rank deformation of large Wigner matrices: convergence and non-universality of the fluctuations, *Annals of Probability* **37**(1), pp. 147.

6. Couillet, R., Silverstein, J.W., Bai, Z.D., and Debbah, M. "Eigen-inference for Energy Estimation of Multiple Sources" to appear in *IEEE Transactions on Information Theory.*

7. Dozier, R.B. and Silverstein, J. W. (2007a). On the empirical distribution of eigenvalues of large dimensional information-plus-noise type matrices', *Journal of Multivariate Analysis* **98**(4) (2007), pp. 678-694.

8. Dozier, R.B. and Silverstein, J. W. (2007b). Analysis of the limiting spectral distribution of large dimensional information-plus-noise type matrices. *Journal of Multivariate Analysis* **98**(6) (2007), pp. 1099-1122.

9. Haagerup, U., Schultz, H., and Thorbjørnsen, S. (2006). A random matrix approach to the lack of projections in $C^*_{\mathrm{red}}(\mathbb{F}_2)$. *Advances in Mathematics* **204**, pp. 1-83.

10. Horn, R.A. and Johnson, C.R. (1985). *Matrix Analysis.* Cambridge University Press, Cambridge.

11. Shohat, J. A. and Tamarkin, J. D. (1970) . *The Problem of Moments.* American Mathematical Society, Providence.

12. Titchmarsh, E.C. (1939). The Theory of Functions, Second Edition. Oxford University Press, London.

13. Vallet, P., Loubaton, P., and Mestre, X. Improved Subspace Estimation for Multivariate Observations of High Dimension: the Deterministic Signals Case. (submitted)

14. Yang, Z.P. and Feng, X.X. (2002). On the trace inequality for products of Hermitian matrix power, *Journal of Inequalities in Pure and Applied Mathematics* **3**(5), Article 78.